

博士論文

日本語事実性解析に関する研究

成田 和弥

2016年1月20日

東北大学 大学院
情報科学研究科 システム情報科学専攻

本論文は東北大学 大学院情報科学研究科 システム情報科学専攻に
博士 (情報科学) 授与の要件として提出した博士論文である。

成田 和弥

審査委員：

乾 健太郎 教授 (主指導教員)

篠原 歩 教授 (副指導教員)

木下 哲男 教授 (副指導教員)

岡崎 直観 准教授 (副指導教員)

日本語事実性解析に関する研究*

成田 和弥

内容梗概

事実性は、文中の事象の成否について、著者や登場人物の判断を表す情報である。事実性解析には、機能表現や、文節境界を越えて事実性に影響を与える語とそのスコープなどの4種類の問題が含まれており、性能の向上が容易ではない。本研究では、事実性解析の課題分析を行うために、機能表現のみを用いたルールベースの事実性解析器を構築し、1,533文に含まれる3,734事象に適用した結果の誤りを分析した。このとき全ての事象表現について、付随する機能表現に対して人手で意味ラベルを付与した。その結果、主事象の事実性解析については、機能表現の意味ラベルが正しく解析できれば、現在の意味ラベルの体系と本研究で用いた単純な規則だけでも、90%に近い正解率が得られることがわかった。従属事象の事実性解析では、後続する述語やスコープといった従属事象特有の誤りが多く見られた。それらの要素についてさらなる分析を行い、今後の事実性解析の指針を示した。そして、そのうちの一つの課題であるスコープ解析について、事実性解析への応用に適合する問題設計を行い、設計したスコープ情報を付与したコーパスを新たに構築し、機械学習に基づくスコープ解析器で8割の正解率を達成した。

キーワード

事実性，モダリティ，機能表現，スコープ

*東北大学 大学院情報科学研究科 システム情報科学専攻 博士論文, 2016年1月20日.

A Study on Japanese Factuality Analysis*

Kazuya Narita

Abstract

Event factuality is information pertaining to whether events mentioned in the natural language correspond to either actual events that have occurred in the real world or events that are of uncertain interpretation. In factuality analysis, sufficient performance is yet to be achieved because of the complexity of issues such as functional expression and linguistic scope. This paper discusses the issues involved in factuality analysis by analyzing errors when applying a rule-based system to 3,734 events in 1,533 sentences. We annotate functional expression labels for all events. In the main events, the factuality analyzer, consisting of simple functional expression rules, achieves approximately 90% accuracy if correct functional expression labels are provided. In subordinate events, we found many errors specific to subordinate events, such as errors caused by predicates and linguistic scopes. We provide guidelines for factuality analysis through additional discussion regarding predicates and linguistic scope. For improvement of factuality analysis, we designed linguistic scope analysis, which is one of the issues involved in factuality analysis. We constructed the corpus for scope analysis and achieved approximately 80% accuracy by scope analyzer based on machine learning.

Keywords:

event factuality, modality, functional expressions, linguistic scope

*Doctor's Thesis, System Information Sciences, Graduate School of Information Sciences, Tohoku University, January 20, 2016.

目次

| | | |
|-------|----------------------------|----|
| 1 | 序論 | 1 |
| 2 | 関連研究 | 5 |
| 2.1 | 言語学におけるモダリティ分類 | 5 |
| 2.2 | タグ体系およびコーパス構築に関する研究 | 9 |
| 2.3 | 解析および課題分析に関する研究 | 14 |
| 2.4 | 生物医学分野における研究 | 15 |
| 2.5 | 事実性に影響を与える要素に関する研究 | 16 |
| 3 | 事実性解析の課題分析のための実験環境の構築 | 19 |
| 3.1 | 問題設定 | 19 |
| 3.2 | 事実性解析に関わる言語要素 | 20 |
| 3.2.1 | 機能表現 | 21 |
| 3.2.2 | 述語周辺の副詞 | 23 |
| 3.2.3 | 文節境界を越えて事実性に影響を与える語とそのスコープ | 23 |
| 3.3 | 課題分析の方針 | 24 |
| 3.4 | 事実性ラベル付与コーパス | 26 |
| 3.5 | 機能表現意味ラベル付与コーパス | 27 |
| 3.5.1 | 機能表現意味体系の設計 | 28 |
| 3.5.2 | 機能表現意味ラベル付与コーパスの構築 | 28 |
| 3.5.3 | 機能表現解析 | 33 |
| 3.6 | 誤り分析に用いる事実性解析モデル | 33 |
| 4 | 事実性解析実験に基づく課題分析 | 40 |
| 4.1 | 主事象に対する事実性解析 | 40 |
| 4.2 | 従属事象における事実性解析 | 45 |
| 4.2.1 | 事象参照表現に後続する述語に関する分析 | 48 |
| 4.2.2 | 事象間の接続表現に基づくスコープに関する分析 | 51 |

| | | |
|-----|----------------------|----|
| 5 | 否定・推量・疑問のスコープとその自動解析 | 59 |
| 5.1 | スコープの付与および事例分析 | 60 |
| 5.2 | 機械学習に基づくスコープ自動解析 | 62 |
| 5.3 | 評価実験 | 63 |
| 5.4 | スコープ解析のまとめ | 67 |
| 6 | 結論 | 69 |
| | 謝辞 | 71 |
| | 参考文献 | 72 |
| | 発表文献一覧 | 78 |

目 次

| | | |
|---|--|----|
| 1 | 事実性に関わる言語要素の構造: 矢印は, 要素が事象表現の事実性に影響することを示す | 21 |
| 2 | 事実性解析課題の切り分け | 25 |
| 3 | スコープ解析の入力例: 係り受け解析結果に機能表現ラベルが付与されたもの | 62 |

表 目 次

| | | |
|----|--|----|
| 1 | Saurí and Pustejovsky による事実性 ([NA] は利用不可値を示す) | 10 |
| 2 | 確信度と肯否極性の組み合わせによる事実性のラベル；下段は拡張モダリティタグ付与コーパスの真偽判断タグとの対応 | 20 |
| 3 | コーパス中の事実性の分布 | 27 |
| 4 | 定義した機能表現意味ラベルの一覧 | 28 |
| 5 | 意味ラベルの付与例 | 30 |
| 6 | 機能表現意味ラベルコーパスの統計情報 | 30 |
| 7 | 作業員間一致率の κ 値 | 31 |
| 8 | 機能表現意味ラベルの作業員間一致率 | 32 |
| 9 | 機能表現解析における学習素性の一覧および素性抽出例 | 34 |
| 10 | 機能表現解析器の評価結果 | 34 |
| 11 | 更新ルールと意味ラベルの対応 | 37 |
| 12 | 機械学習モデルで用いた素性一覧および (16) における素性抽出例 | 39 |
| 13 | 主事象に対する事実性解析の評価 | 41 |
| 14 | 主事象に対する事実性解析の各軸ごとの評価 | 41 |
| 15 | 誤りの種類の分布 (カッコ内は、事実性のアノテーション誤りを除いた部分での誤りの割合) | 42 |
| 16 | 従属事象に対する事実性解析の評価 | 45 |
| 17 | 従属事象に対する事実性解析の各軸ごとの評価 | 45 |
| 18 | 誤りの種類の分布 (カッコ内は、事実性のアノテーション誤りを除いた部分での誤りの割合) | 46 |
| 19 | 事象選択述語辞書の記述例 | 49 |
| 20 | 誤り事例における事象選択述語；カッコ内は 25 事例中の延べ数を示す | 50 |
| 21 | 誤り事例における接続表現の分類 | 57 |
| 23 | 主事象と最も近い従属事象との間の接続表現の分類 | 57 |
| 22 | ランダムに抽出した 140 文中の従属事象の分布 | 57 |
| 24 | スコープのアノテーションによる事実性解析性能 | 58 |

| | | |
|----|---|----|
| 25 | スコープのアノテーションによる事実性解析結果の変化 | 58 |
| 26 | 付与されたスコープの割合 | 62 |
| 27 | スコープ解析の素性 | 63 |
| 28 | 訓練用データと評価用データにおけるスコープの割合 | 64 |
| 29 | スコープの自動解析結果 (機能表現の意味ラベルごとに分割した 場合とそうでない場合との比較); 太字は性能が高いもの | 64 |
| 30 | スコープの自動解析結果 (アプレーションテスト); *は全素性利 用時と比較して性能の低下が見られたもの | 65 |
| 31 | スコープの自動解析結果 (素性ごとの比較) | 65 |
| 32 | スコープの自動解析結果 (オープンテスト); ベースラインは素性 aのみを用いて学習したもの | 66 |
| 33 | オープンテストにおける Confusion Matrix ; 太字は正解事例 | 67 |

1 序論

近年，ブログ等の個人が自由に情報を発信できる環境の爆発的な普及に伴い，膨大なテキスト情報が Web 上に加速度的に蓄積され，利用できるようになってきている．これらの情報を整理し，そこから有益な情報を得るためには，「誰が」「いつ」「どこで」「何を」といった情報を認識するだけでなく，文に記述されている事象が，実際に起こったことなのかそうでないことなのかという情報を解析する必要がある．我々はこのような，文の著者や文中の登場人物による事象の成否に対する判断情報を事実性と呼ぶ．

- (1) a. 商品 A を使い 始めた。
- b. 商品 A を使う のは簡単ではなかった。
- c. 商品 A を使って みたい。
- d. 商品 A を使っ ているわけではない。
- e. 商品 A を使っ ているはずだ。

(1) に示す例は，いずれも「商品 A を使う」という事象が含まれるが，その事実性は異なる．(1a) と (1b) は，事象が成立していると解釈できる一方で，(1c) と (1d) は，事象は成立していないと解釈できる．さらに (1e) は，事象の成立を推量していると解釈できる．評判分析などの文脈で，商品 A を使っているユーザの情報のみを抽出したい場合，(1) に示した全ての文に対して，「商品 A を使う」と照合するだけでは，(1c) や (1d) といった，商品 A を実際には使っていないユーザの情報まで抽出されてしまう．そこで事実性解析を用いると，(1a) や (1b) が実際に商品 A を使っており，(1c) や (1d) が使っていない，(1e) は使っていない可能性がある，ということを区別することができる．事実性解析は，評判分析だけでなく，含意関係認識や知識獲得といった課題に対しても重要な技術である [1, 2, 3] ．

先行研究では，事実性だけでなく，時制などの関連情報についても，付与基準が議論されるとともに，コーパス構築が進められてきた [4, 5, 6, 7] ．日本語を対象とした事実性解析の研究は少なく，述部（本研究の事象表現に相当）に続く表

現に着目したルールベースの解析 [8] や機械学習に基づく解析器 [9] など、わずかな報告があるに留まっている。前者はその性能は報告されていないが、後者の解析性能は、9種類の事実性ラベルの分類性能がマクロ F 値で 48% であり、実用上十分とはいえない。

事実性解析の性能向上が困難である理由の一つは、述語に後続する機能表現の多様性にある。詳しくは 3 章で述べるが、例えば「<<使わ>>ない」「<<使う>>わけない」「<<使わ>>ねえ」「<<使う>>もんか」のように、事象が成立しない（あるいは、成立していない）ことを示す機能表現（下線部）が多々ある。ここで、<<> は事象表現、即ち事象の中心的な述語を示す記号であり、事象の事実性は事象表現に割り当てられると定義する。機能表現以外に、「<<使う>>のをやめた」のように、文節境界を越えて事象の不成立を示唆する述語（下線部）の存在もあり、さらにこれらの要素の組み合わせが、事実性解析の性能向上を阻んでいる。

本研究の最終目標は事実性解析の性能向上である。そのためにまず、機能表現や文節境界を越えて事実性に影響する述語などの、事実性に影響を与える言語要素を、可能な限り切り分けて課題を分析する。本研究では、節の違いによる、事実性に影響を与える言語要素の違いに着目し、以下の 2 つの仮説を構築する。

仮説 1 主節の事象の事実性を決定する要因は、主節の述語に続く機能表現と主節の述語を修飾する副詞が支配的である。

仮説 2 従属節の事象は、節内の機能表現、副詞に加えて、主節を含む他の節の述語および機能表現の影響を受け、それらの相互作用によって事実性が決定される。

これらの仮説に基づくと、機能表現に基づく事実性解析器を構築し、主節と従属節に分けて誤り分析をすることで、複雑に組み合わせる要素を切り分け、どの要素に取り組むことが重要であるかを明らかにすることができる。同時に、これらの仮説が真であるかを検証することができる。

具体的には、まず、事実性が付与されたコーパス [5] について、述部に後続する機能表現に意味ラベルを付与する。結果的に、1,533 文に含まれる 3,734 事象に対して機能表現に意味ラベルを付与した。そして、機能表現の意味ラベルに基づい

て、決定的に事実性を解析するモデルを提案するとともに、その誤り分析によって事実性解析の課題分析を行う。本研究では、構築した事実性解析器を 3,734 事象 (1,533 文) に適用し、1,533 個の主節の事象と 2,201 個の従属節の事象とに分割し、それぞれについて評価および誤り分析を行った。

主節の事象については、事象表現に後続する形態素は機能表現のみで構成されるため、誤り分析の結果によって、機能表現または副詞に関する問題が大部分であることが示されれば、仮説 1 は正しいと判断できる。実験の結果、機能表現の意味ラベルが正しく解析できれば、事実性解析の正解率は約 90% となることが分かった。誤り分析の結果、機能表現の問題を除いた残りの半数は副詞に起因するものであることが分かった。以上のことから、仮説 1 は正しいと判断できた。

一方で、従属節の事象については、事象表現に後続する形態素として、付随する機能表現以外に、主節を含む他の節が存在するため、誤り分析の結果によって、どのような作用によって従属節の事象の事実性が決定されるかを示すことで、仮説 2 を検証する。従属節の事象を評価したところ、主節の事象に比べて事実性の正解率は低くなった。誤り分析の結果、従属節でのみ考慮すべき要素は大きく二つあり、文節境界を越えて事実性に影響を与える述語と、文末側にある他の節の事象に含まれる機能表現の影響である。前者は、既存の辞書のカバレッジを調査した結果、これを利用することで誤りの一部を解消できるものの、さらなる拡充が必要であることが分かった。後者は、問題となるケースは多様ではなく、事象間の接続表現によってある程度決定できることを明らかにした。決定不可能な接続表現については、隣接する事象の機能表現が及ぼす範囲 (スコープ) を同定する問題を設計し、機械学習に基づく手法を提案した。その際に、1,533 文とは別に 3,589 文を用意し、従属節の事象が主節の事象の機能表現の影響を受けるか、すなわちスコープ内にあるかを付与した。

本論文の貢献は、大きく以下の 3 点である。

1. 事実性が付与されたコーパスである拡張モダリティタグ付与コーパス [5] に対して、述部に続く機能表現に対する意味ラベルを設計し、それを付与することで、機能表現と事実性の情報を重層的に付与したコーパスを構築した。
2. 機能表現の意味的抽象化および事実性の構成性 (相互作用モデル) に基づ

く事実性解析手法を提案するとともに，その誤り分析によって事実性解析の課題分析を行った．その結果，主節の事象，従属節の事象ともに機能表現解析が主要な問題であること，また，従属節の事象においてはスコープ解析が問題であるものの，問題になるケースは，これまでの想定に比べて限定的であることを明らかにした．

3. 事実性解析を主眼に置いた日本語のスコープ解析課題を設計し，スコープ解析が問題となるケースに限定して3,589文に対してスコープを付与し，自動解析器を構築した．

本論文の構成は以下の通りである．2章では事実性解析およびモダリティ解析の関連研究について述べる．3章では誤りに分析に基づく事実性解析の課題分析の方針について述べる．4章では，エラー分析に基づく課題分析結果について述べる．5章では，従属事象の課題として重要なスコープ解析について，問題設計と解析器の構築について述べる．6章でまとめる．

2 関連研究

事実性に大きく関連する概念として、態度表明者の主観的な態度（モダリティ）、および、肯定/否定があげられる。本章では、まず言語学におけるモダリティやその周辺要素に関する言及について述べる。次に、自然言語処理分野における、事実性に関連する情報を付与するためのタグ体系・コーパス構築、および、その解析手法、そして機能表現やスコープなどの事実性に影響を与える要素に関してまとめる。

2.1 言語学におけるモダリティ分類

事実性やモダリティ、肯定/否定などの分類に関しては、言語学においてこれまで様々な議論が行われてきている。言語学において、文全体の意味は、事態を表す意味領域、および、話し手の態度を表す意味領域に分けられ、前者は「命題」、後者は「モダリティ」と呼ばれる¹[10, 11, 12]。

- (2) a. ねえ、どうやら昨夜激しく雪が降ったようだよ。
- b. [ねえ、どうやら [昨夜激しく雪が降った] ようだよ]
- c. [ねえ [どうやら—ようだ] よ]

例えば、(2a)の文は、(2b)のような意味領域からなり、「昨夜激しく雪が降った」が事態を表す命題の領域、「ねえ、どうやら—ようだよ」が態度を表すモダリティの領域である。さらに、このモダリティの領域は、文の意味的階層構造という観点から、事態に対する判断を表す領域、および、表現・伝達を表す領域に分けられ、前者を「判断のモダリティ」、後者を「発話のモダリティ」と呼ぶ。(2a)におけるモダリティの領域は、(2c)のような2つの領域に分けられ、「どうやら—ようだ」の部分が判断のモダリティの領域を、「ねえ—よ」の部分が発話のモダリティの領域を表している。

¹Fillmore (1968) [10] は、文が proposition (命題) と modality (モダリティ) の2つで構成される、という見方をした代表的な研究であり、「Sentence = Proposition + Modality」という書き換え規則を導入している。

益岡 (2007) [11] は、モダリティを以下の7種類のカテゴリーに分類している。

真偽判断のモダリティ 事態が成り立つかどうかの真偽性(断定か、非断定か)を表す。非断定の判断は、断定こそできないものの何らかの判断は下すという「定判断」と、疑問文など、真偽の判断がまったく下せない「不定判断」に大別される。さらに、定判断には、「だろう」「～(よ)う」などの表現によって、真であるとの確信が持てなかったり、聞き手との関係で断定を差し控えたりする場合を表す「断定保留」、「かもしれない」(可能性)、「にちがない」(必然性)といった、確からしさの程度を表す「蓋然性判断」、「ようだ」「らしい」など、ある証拠に基づいて推定を行うことを表す「証拠性判断」、「はずだ」といった、推論から得られる当然の帰結を表す「当然性判断」という4つの下位類に分類される

価値判断のモダリティ 事態が是認されるかどうかの妥当性(現実像か、理想像か)を表す。理想像を表す形式は、「べきだ」「ほうがよい」などの「適当」、「なければいけない」「しかない」などの「必要」、「てもよい」「てはいけない」などの「容認・非容認」といった意味に分類される

発話類型のモダリティ 文法的な観点から見た、表現・伝達の類型的な機能を表す。話し手の認識を表す「演述型」、話し手の内面にある感情や意志を表す「情意型」、聞き手に判断を求める、あるいは、話し手の判断が定まらない状況を表す「疑問型」、聞き手に何らかの対応を求める「要求型」、話し手の感情の発露を表す「感嘆型」のいずれかの態度を表す

丁寧さのモダリティ 対話文において、聞き手に対する丁寧さの有無を表す。「です」「ます」といった表現がこれに該当する

対話態度のモダリティ 聞き手に対する情報提示の調整を表す。「ね」「よ」「よね」などの終助詞がこれに該当する

説明のモダリティ 種々の説明を表す。「のだ」「わけだ」「ものだ」などの表現があり、各表現によって用法が変化する。例えば「のだ」には、「叙述様式説

明」「事情説明」「帰結説明」「実情説明」「当為内容説明」といった用法が存在する

評価のモダリティ 述部以外の付加部に出現するモダリティ要素であり、事態に対する評価を表す。「あいにく」「さいわい」「親切にも」のように、通常文頭に現れる

「判断のモダリティ」の下位カテゴリとして、「真偽判断のモダリティ」および「価値判断のモダリティ」、「発話のモダリティ」の下位カテゴリとして、「発話類型のモダリティ」および「丁寧さのモダリティ」、「対話態度のモダリティ」を認定し、さらに、特殊なモダリティとして、判断のモダリティの領域と発話のモダリティの領域の両方に関係する「説明のモダリティ」、および、述語以外の部分に現れる要素である「評価のモダリティ」を認定している。

文献 [12] では、モダリティを以下の4つのタイプに大別している。

文の伝達的な表し分けを表すモダリティ 聞き手に情報を伝達する「叙述のモダリティ」、聞き手から情報を聞き出そうとする機能をもつ「疑問のモダリティ」、話し手の行為の実行を表す「意志のモダリティ」、話し手の行為の実行を前提として、聞き手に行為の実行を求める「勧誘のモダリティ」、命令など、聞き手に行為の実行を求める「行為要求のモダリティ」、物事に触れて引き起こされる話し手の感動の気持ちを表す「感嘆のモダリティ」といったタイプに分類される。このモダリティ全体で、益岡 [11] の「表現類型のモダリティ」に相当する

命題が表す事態のとらえ方を表すモダリティ 命題によって表される事態に対する、必要、不必要、あるいは許容できる、できないといった、話し手の評価的なとらえ方を表す「評価のモダリティ」、断定や推量、可能性や必然性、推定や伝聞など、事態に対する話し手の認識的なとらえ方を表す「認識のモダリティ」といったタイプに分類される。「評価のモダリティ」は益岡 [11] の「価値判断のモダリティ」²、「認識のモダリティ」は「真偽判断のモダリ

²益岡 (2007) [11] にも「評価のモダリティ」という分類が存在するが、文献 [12] における「評価のモダリティ」とは別のものである。

ティ」に相当する

先行文脈と文との関係付けを表すモダリティ その文を先行文脈と関係があるものとして示すことによって、先行文脈の内容が聞き手に理解させやすくする働きをもつ。

(3) 遅れてすみません。渋滞していたんです。

(3)の文では、先行文脈で表されている、話し手が遅れたという事実に対して、道路が渋滞していたという後続文で表されている事実が「のだ」によって関係付けられて示されている。このモダリティは、益岡 [11] の「説明のモダリティ」に相当する

聞き手に対する伝え方を表すモダリティ 聞き手に対してその文を通常スタイルで伝えるか、丁寧なスタイルで伝えるかという、スタイルの選択に関わる「丁寧さのモダリティ」、話し手の認識状態を示したり、聞き手に伝えるにあたっての微調整をしたりする「伝達態度のモダリティ」といったタイプに分類される。それぞれ、益岡 [11] の「丁寧さのモダリティ」、「対話態度のモダリティ」に相当する

しかしながら、この分類では、益岡 [11] の「評価のモダリティ」に相当するものが考慮されていない。以上のように、我々の調べた限りでは、用語も含めて、完全に統一された見解は存在しないようであるが、概念としてはおおむね対応がみられる。

この中で、事象の真偽に対する書き手の確信度を表した「真偽判断（評価）のモダリティ」は、本研究における事実性と非常に近い概念である。また、文献 [13] によると、事態の成立を表すことを肯定といい、事態の不成立を表すことを否定という。以上のことから、本研究における事実性は、言語学における「真偽判断のモダリティ」と、肯定または否定の組み合わせに相当する。

2.2 タグ体系およびコーパス構築に関する研究

事実性およびその周辺情報を付与するためのタグ体系およびコーパス構築の関連研究として、Prasadらによる Penn Discourse TreeBank [14] における attribution タグや、Saurí and Pustejovsky [15, 4] による FactBank、松吉ら [5, 16] による拡張モダリティタグ付与コーパスなどがある。

Prasadら (2008) [14] は、Penn Discourse TreeBank (PDTB) の談話関係とその項に対して、情報の発信源とその情報が事実であるかどうかを判断するために、attribution という属性タグを付与している。attribution タグは以下の4種類の要素からなる。

Source 情報の発信源や判断している主体を表し、テキストの著者 ("Wr")、テキストにより導入される特定の動作主 ("Ot")、任意の個人や組織 ("Arb")、前の関係を継承 ("Inh") のいずれかの値をとる

Type 対象が事実である度合いを推論するための意味クラスであり、主張 ("Comm")、信念 ("PAtt")、事実 ("Ftv")、成り行き ("Ctrl") のいずれかに分類される

Scopal Polarity 作用の及ぶ範囲(スコープ)を考慮した否定の有無 ("Neg", "Null")

Determinacy 上記3つの要素のいずれかが、より広い文脈で非決定的になりうるかどうか ("Indet", "Null")

彼らは、Typeにおいて、動詞の意味クラスのみを用いて、事実性を表している。しかし、助詞の"may"や、副詞の"probably"のように、動詞以外にも事実性に影響する要素があるため、事実性を動詞の意味クラスのみで表すのは不十分である。

Saurí and Pustejovsky (2009) [15, 4] は、事象を対象とし、以下の2つ組のタグによって事実性を定義した。

modality 事実らしさに対する態度表明者の確信度。CT (Certain)、PR (Probable)、PS (Possible)、U (Underspecified) の4種類で表す

polarity 事象に対する確信の方向。+ (positive)、- (negative)、u (underspecified) の3種類で表す

表 1: Saurí and Pustejovsky による事実性 ([NA] は利用不可値を示す)

| 確信度 \ 肯否極性 | positive (+) | negative (-) | underspecified (u) |
|--------------------|----------------|--------------------|----------------------------------|
| Certain (CT) | fact (CT+) | counterfact (CT-) | certain but unknown output (CTu) |
| Probable (PR) | probable (PR+) | not probable (PR-) | [NA] |
| Possible (PS) | possible (PS+) | not possible (PS-) | [NA] |
| Underspecified (U) | [NA] | [NA] | unknown or uncommitted (Uu) |

これらの組み合わせによる事実性の定義を表 1 に示す．例えば，事象が実際に起こったことである，ということ CT+ と表す．そして，事象とその時間情報や，事象間の時間的順序関係が付与された TimeML [17] の上に，確信度と肯否極性を態度表明者 (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．以下に Saurí and Pustejovsky [4] による FactBank のアノテーション例を示す．

(4) He does not think_{e₀} she followed_{e₁} the rules.

$f(e_0, \text{author}) = \text{CT-}$

$f(e_1, \text{author}) = \text{Uu}$

$f(e_1, \text{he_author}) = \text{PR-}$

$f(e, s)$ は態度表明者 s から見た事象 e の事実性を示している．この文では，著者から見た e_0 (think) の事実性が CT-，著者から見た e_1 (followed) の事実性が Uu であることが付与されるとともに，文中の登場人物 he から見た e_1 (followed) の事実性を著者は PR- と判断している，ということが付与されている．de Marneffe *et al.* (2012) [18] は，PR+ と PS-，PS+ と PR- をそれぞれ区別せず，同一のラベルとして取り扱い，5 種類のラベル体系による評価を行っている．

松吉ら (2010) [5, 16] 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，それを現代日本語書き言葉均衡コーパス (BCCWJ)³ の各事象に付与したコーパ

³http://www.ninjal.ac.jp/corpus_center/bccwj/

スを構築した。

態度表明者 態度を表明している人物「wr:筆者」(態度表明者が書き手である)や「wr:筆者_arb:不特定」(態度表明者が不特定の個人や集団である,と書き手が述べている),「wr:筆者_l:太郎」(態度表明者が太郎である,と書き手が述べている)のように記述する

相対時 態度表明時に対する相対的な時制「未来」または「非未来」(過去,現在,脱時間的のいずれか)の2種類で表現する

仮想 文章に記述される情報が事実であるのか,それとも,単なる仮想的な話であるのかを示す「条件」(事象が条件として仮想的に述べられている),「帰結」(事象が仮想的な条件の帰結として述べられている),「0」(2つのいずれでもない)の3種類で表現する

態度 命題に対する態度表明者の主観的な態度であり,益岡(2007)[11]における「発話類型のモダリティ」に相当する「叙述」(事象が表す内容や態度表明者の判断などを情報の受け手に伝える),「意志」(態度表明者が,自分自身の行為の実行,もしくは,非実行を決定),「欲求」(態度表明者が,自分自身の行為の実行,もしくは,非実行を望んでいる),「働きかけ-直接」(態度表明者が,直接,相手に対して,行為の実行,もしくは,非実行を求める),「働きかけ-間接」(態度表明者が,間接的な表現手段を用いて,行為の実行,もしくは,非実行を相手に求める),「働きかけ-勧誘」(態度表明者が,態度表明者の行為を前提として,行為の実行,もしくは,非実行を相手に誘いかける),「許可」(態度表明者が,事象の成立,もしくは,不成立を許容する),「問いかけ」(態度表明者にとって不明なことがあるために,その事象に対して態度表明者の判断が成り立たない)の8種類で表現する

真偽判断 事実の真偽に対する態度表明者の確信度であり,Sauriら(2009)[4, 15]の事実性に相当するといえる「成立」(肯定の断定),高確率(肯定の推量),低確率(否定の推量),「不成立」(否定の断定),0(詳細不明)の5種類に加え,これらの間を表す真偽の変化を含意するラベルとして,「成立から不

成立」、「高確率から低確率」、「低確率から高確率」、「不成立から成立」の4種類，計9種類で表現する

価値判断 事象成立の望ましさを表す極性情報「ポジティブ」(態度表明者が，事象成立が望ましいと判断している)、「ネガティブ」(態度表明者が，事象成立は望ましくないと判断している)、「0」(事象成立の望ましさについての態度表明者の判断が，文章中に記述されていない)の3種類で表現する

以下にこのコーパスにおけるアノテーション例を示す．

- (5) ソフトを新品で 買お うと 思っているのですが、どこで 買ったら一番 安い
 でしょうか？

| 事象 | 態度表明者 | 相対時 | 仮想 | 態度 | 真偽判断 | 価値判断 |
|----|-------|-----|----|------|------|-------|
| 買お | wr:筆者 | 未来 | 0 | 意志 | 高確率 | ポジティブ |
| 思っ | wr:筆者 | 非未来 | 0 | 叙述 | 成立 | 0 |
| 買っ | wr:筆者 | 未来 | 条件 | 問いかけ | 0 | 0 |
| 安い | wr:筆者 | 未来 | 帰結 | 問いかけ | 0 | 0 |

(5) では，4つの事象に対して，6項目からなる拡張モダリティタグがそれぞれ付与されている．例えば「買お」という事象では，ソフトを新品で買うつもりである，という著者のポジティブな意志が示されており，未来において高確率で買うだろうと著者は考えている，ということ表現したラベルとなっている．彼らのタグ体系の内，<真偽判断>は Saurí *et al.* [4] の事実性に相当している．拡張モダリティタグ体系を用いた解析では，項目間の依存関係を考慮することが可能であるが，それ故に処理が複雑化してしまうという問題がある．

川添ら (2011) [6, 7] は，テキストに現れる事実とそれ以外の情報との区別，また推量や仮定など間に見られる確実性の差を自動的に識別することを目指している．そのために，以下のように「確実性」に影響を与える言語表現を分析・分類し，それに従ってそれらの言語表現およびその影響を与える範囲(スコープ)をアノテーションした MCN コーパスを構築している．

様相表現 「ようだ」や「だろう」のような，文の内容に対する書き手/語り手の認識・判断を表す表現であり，益岡 (2007) [11] で「真偽判断のモダリティ」

と呼ばれているものに対応する．叙実表現，証拠推量表現，認識的推量表現，他人の認識を表す表現，不定判断・疑問表現，比況表現，反叙実表現に分類される

否定表現 「わけではない」のような通常の否定と「～（というのは）正しくない」のようなメタ否定 [19] を区別している．通常否定の解釈は命題を否定する解釈だが，メタ否定の解釈は先行発話そのものが不適切であると異を唱える解釈である

条件表現 「たら」や「なら」，「としても」のような条件を表す表現であり，事実的条件表現，予測的条件表現，認識的条件表現，一般的条件表現，半事実条件表現に分類される

この「確実性」は「テキストの書き手が判断する，命題の内容が真である確率」という意味で定義されており，我々の事実性と対応している．また，様相表現，否定表現，条件表現のアノテーションにあたっては，テキスト上での出現について，その意味や用法を特定し，適切なラベルを選択するという，曖昧性解消が課題となる．田中ら [20, 21, 22] は，これらの表現に関して意味，用法を詳細に分類しており，その曖昧性解消のための言語学的テストの設計も行っている．田中らの用法分類は非常に詳細であり，それらを構成的に用いることで事実性の解析に利用できると考えられるが，その細かさ故に自動分類が難しく，また現在利用可能なデータ量も多くない．

本研究では，de Marneffe *et al.* [18] の枠組みに基づいて事実性のラベルを定義する．この枠組みを利用することで，事実性を確信度と肯否極性の2軸に分けることができるため，問題の分析がしやすくなると考えた．誤り分析には，松吉ら [5, 16] の拡張モダリティタグ付与コーパスを用いる．事実性解析における課題分析をする上で十分な量であり，一般に利用可能なコーパスは他にないため，拡張モダリティタグのうち〈真偽判断〉を事実性の正解として利用し，事実性解析の誤り分析を行う．

2.3 解析および課題分析に関する研究

事実性は、機械学習に基づく手法や、人手で構築した語彙的・統語的な知識を利用したパターンベースの手法などを用いて解析が行われているが、その性能および課題分析は十分でない。

原ら (2008) [23] は、事象の事実性情報を、<時間情報(極性を含む)、モダリティ、モダリティの時間情報(極性を含む)>の3種類で表現し、SVMを学習器に用いた解析手法を提案している。素性は、予測対象の文節、その前後の文節、文全体を区別した上で、品詞と原型を組み合わせたものを用いた。Inuiら (2008) [24] は、原らの提案するタグ体系を整理統合し、条件付き確率場を学習器として用いた解析手法を提案した。実験の結果、SVMを用いるよりも、タグ間の依存関係を考慮できる条件付き確率場を用いたほうが、精度が高いことが示されている。江口ら (2010) [9] は、拡張モダリティタグ体系における項目間には強い依存関係があり、また、同じ文に存在する複数の事象間にも依存関係があることに着目した。そこで彼らは、項目間、および事象間の依存関係を考慮できる条件付き確率場を用いた拡張モダリティ解析システムを構築した。事実性に関連の深い<真偽判断>には9種類のラベルが存在するが、そのマクロF値で48%の性能を示している。さまざまな枠組みによって事実性の解析が行われているが、いまだ十分な性能は達成できておらず、その課題を分析する余地が多分に残されている。

モダリティ解析における課題分析としては、松吉ら [16] が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分析を試作し、その中の1つの項目である<態度>に着目した誤り分析を行っている。彼らは語義曖昧性解消や連体節内の述語に及ぼす影響の解明、節間の意味的關係の認識などが、<態度>に関するモダリティ解析の精度向上に向けた課題であることを述べている。事実性に、より直接的に関連する<真偽判断>の誤り分析でも同様の結果が得られるかどうかは明らかではない。

英語においては、Saurí and Pustejovsky [25] や de Marneffe *et al.* [18] が事実性の解析に取り組んでいる。Saurí and Pustejovsky [25] は、事象の成立に影響を与える手がかり表現を利用し、態度表明者ごとに、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、パターンベースの決定的アルゴ

リズムを提案した．例えば *not* があれば肯否極性を反転させる，*may* があれば確信度を下げる，といったルールに基づいて解析を行い，F 値でマクロ平均 70%，マイクロ平均 80%の性能を実現している．誤り分析の結果，ルールのカバレッジや表現の曖昧性が大きな問題であることを報告している．Saurí and Pustejovsky [25] のアノテーション基準では，事実性は可能な限り客観的に判断される．一方で，de Marneffe *et al.* [18] は，主観的な判断の自動推定に取り組んだ．主観的な判断とは，例えば態度表明者の社会的な信頼性によるものである．信頼に足る組織が表明した事象の事実性は，CT+ にバイアスがかかるが，表明者が不明な場合は事象の事実性は，CT- にバイアスがかかる．彼女らは，FactBank 中の各事象に対して，10 名ずつアノテーションを行い，その分布を最大エントロピーモデルによって推定した．解析性能は，多数がアノテートしたラベルを正解とした場合に，F 値でマクロ平均 70%，マイクロ平均 83%の性能をあげられている．

本研究では，日本語事実性解析の課題に関して議論するために，機能表現に基づき，決定的に事実性を解析するルールベースのモデルを構築し，誤り分析を行う．ここでルールベースモデルを用いる理由としては，機械学習に基づく手法と比べ，出力結果がどのような要素に基づいて選択されたかがわかりやすく，本研究の目的とする，日本語事実性解析における課題の分析に対して適当であると判断したためである．また，事実性に影響を与える要素はさまざま存在しており，いろいろな要素を複合的に加味したモデルが提案されてきている．しかしながら，どの要素がどの程度事実性に影響を与えるのか，という分析は十分に行われていない．そこで，事実性に影響を与える要素を切り分けることにより，事実性解析における各要素の重要性を議論し，課題の分析を行う．

2.4 生物医学分野における研究

生物医学分野においても，近年このような研究は，盛んに行われている [26, 27, 28]．生物医学分野では，例えば，感染症情報など緊急の判断が必要とされる情報の取得の際に，不要な情報を取り除いて効率的に情報を取り出すことや，情報の確実性・信憑性判断に関わる人的コストを減らせる，などの応用が考えられる [6]．

Light ら (2004) [26] は，生物医学分野の論文のアブストラクト内の，推測を表

す文を人手によってアノテーションしたコーパスを構築した．そして，*suggest, potential, likely, may* などの 14 のキーワードの有無により，推測を表す文かどうかを判断した結果が，SVM による学習結果と同等のパフォーマンスを示した，と報告している．しかしながら，日本語においては，言語表現が多様であり，このようなキーワードによる単純な判断は難しいものと思われる．

Medlock ら (2007) [27] も，生物医学分野のテキストを対象に，推測を表す文を人手によってアノテーションしたコーパスを構築した．彼らは，Light らと異なり，推測を表す文に出現する，特徴的な表現を認識することに重きを置き，スコープはその表現を含む文全体とみなしている．そして，特徴的な表現を自動的に認識する教師あり学習の確率モデルを提案した．

生物医学分野のテキストを対象としたコーパスも構築されている．BioScope (2008) [28] は，否定表現，様相表現，そして，それらのスコープをマークアップしたコーパスであり，スコープを特定する研究等に利用されている．

2.5 事実性に影響を与える要素に関する研究

事実性に影響を与える要素としては，機能表現や後続する述語，および，それらの作用する範囲（スコープ）などがある．

(6) a. もう《遅い》から、彼は先に《帰っ》ている だろう。

b. 問題が《発生する》のを 防いだ。

例えば，(6a) の事象「帰る」の事実性は「だろう」という機能表現に影響を受け，(6b) の事象「発生する」の事実性は「防いだ」という述語に影響を受けている．また，(6a) では「だろう」という機能表現は「帰る」のみに影響を与え，先行する事象「遅い」には影響しない，というように，機能表現や後続する述語の作用する範囲，即ちスコープを特定することも，事実性解析において重要な要素だと考えられる．

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り，このような日本語機能表現の意味に関連した研究が多く進められている．例

えば、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』[29]が利用されている。この辞書は、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、音韻的变化などを網羅的に収録した辞書であり、機能表現の意味として、「対象」や「目的」、「名詞化」など、89 種類のラベルが定義されている。その中には「推量」や「否定」、「疑問」など、事実性に影響を与えるラベルも多数含まれている。また、機能表現の中には表層を見ただけでは判別が難しいものも存在する。

(7) パソコンが <<壊れ>>てしまったかも知れない。

(7) では、事象「壊れる」に対して、「てしまっ」「た」「かも知れない」という機能表現が付随している。「知る」という表現は、機能表現の一部として用いられるだけでなく、述語としても用いられるため、(7) では「かも知れない」で 1 つの機能表現として用いられている、ということ判別する必要がある。このような曖昧性を解消するため、どの部分が機能表現なのかを特定し、その意味を同定する研究も行われている [30, 31, 32]。

事実性に影響を与える述語に関する研究としては、江口ら [9] が構築した、モダリティ解析手がかり表現辞書がある。彼らは「防いだ」のような、拡張モダリティに影響する動詞、形容詞が存在していることに着目した。こうした動詞、形容詞が直前の事象に与える影響を記述した、モダリティ解析のための手がかりを集めた表現辞書を作成し、機械学習による拡張モダリティ解析を行う上で、素性として利用している。このような表現を集めた利用可能なリソースは他に存在しておらず、この辞書を利用することでどの程度事実性解析の性能改善につながるのか、この辞書でどの程度の述語がカバーできているのか、といったことを調査する必要がある。

事実性を決定する上で、否定や推量などのスコープを決定することは重要だと考えられる。否定表現および推量表現のスコープを同定する研究は、近年盛んに行われている。例えば BioScope [28] は、医学・生物学ドメインのテキストを対象に、否定表現、様相表現、そして、それらのスコープをアノテーションしたコーパスであり、このコーパスを用いて Shared Task [33, 34] が開催されるなど、スコープを特定する研究に広く利用されている。日本語においては、川添ら [6, 7] が、テ

キストに現れる事実とそれ以外の情報との区別，また推量や仮定などの間に見られる確実性の差を自動的に識別するため，様相表現，否定表現といった「確実性」に影響を与える言語表現を分析・分類し，それに従ってそれらの言語表現およびそのスコープをアノテーションしたコーパスを構築しているが，それらの定量的な分析を行うまでには至っていない．松吉 [35] は，否定の焦点検出システムを構築するための基盤として，日本語における否定の焦点をテキストにアノテーションする枠組みを提案し，否定の焦点コーパスを構築している．否定の焦点は，否定のスコープの中で特に否定される部分であるため，焦点の検出はスコープの特定と密接に関連している．

このように，事実性に影響を与える要素がいくつか存在しており，これらの要素を複合的に考慮することで事実性を決定できると考えられる．しかしながら，どの要素がどの程度事実性に影響を与えるのか，ということは明確ではない．本研究では，これらの事実性に影響を与える要素を切り分け，事実性解析における各要素の重要性を議論することにより，課題の分析を行う．

3 事実性解析の課題分析のための実験環境の構築

本章では、事実性解析の課題分析を行うために構築した、コーパスおよび事実性解析モデルといった実験環境について述べる。まず、3.1節で、事実性解析の問題設定について述べる。次に、3.2節で、事実性解析に関わる言語要素について概観する。事実性を決定するためには、複数の言語要素が単体だけでなく組み合わせによる影響を考慮する必要がある。3.3節では、組み合わせをどのように切り分けるのか、課題分析の方針について述べる。3.4節では、課題分析のために用いる事実性ラベル付与コーパスについて述べ、3.5節では、機能表現意味ラベルの設計、および、機能表現意味ラベル付与コーパスの構築について述べる。3.6節では、事実性解析の解析モデルについて述べる。

3.1 問題設定

事実性解析は、事象が実際に起こったかを解析する技術ではあるが、真に起こったかどうかを与えられた文のみから判断することは不可能である。例えば「太郎は先に帰ったはずです。」という文において「太郎は帰った」という事象について著者は事実であると推量しているが、真に事実か否かは「太郎」にしか分からず、この文からは「太郎」の判断結果は分からない。また、本研究で用いる事実性が付与されたコーパスには、松吉ら [16] によって、著者および登場人物ごとに判定された事実性が付与されているが、文中の人物によって事実性が判定されている事象は、全体の1割程度にとどまっていることが分かっている。以上を背景として本研究では、事実性を、文中の事象の成否について、著者の判断を表す情報と定義する。ただし、実際には著者の判断も真にはわからないため、著者の判断を読者がどう解釈できるかによって事実性を表す。前述の例では、著者は事象「太郎は帰った」の成立を推量していると読者は解釈するのが自然であろう。

事実性の付与対象となる事象は、松吉ら [5] と同様に、行為、出来事、状態の総称であると定義する。

(8) a. 雨が<<降っ>>_{出来事}たら、バスで<<行き>>_{行為}ます。

表 2: 確信度と肯否極性の組み合わせによる事実性のラベル；下段は拡張モダリティタグ付与コーパスの真偽判断タグとの対応

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

b. <<混雑>>状態していたら、別のところに<<行き>>行きます。

(8) に示す例では、「(雨が)降る」、「(バスで)行く」、「混雑する」、「(別のところに)行く」が全て事象である。アノテーションや解析において、事実性のラベルは<<>>で囲まれた事象表現に付与する。

付与する事実性ラベルは、Saurí and Pustejovsky [4] の体系を一部簡素にした de Marneffe *et al.* [18] によるラベル体系を採用する。本体系は、3種類の確信度と、3種類の肯否極性の2軸に分けて事実性を定義しており、それぞれの軸で評価できることが、課題分析に有効であると考えた。後述する拡張モダリティタグ付与コーパスの真偽判断タグと、事実性ラベルとの対応を表2に示す。

まとめると、本論文で構築する事実性解析器は、文を入力として、形態素解析および係り受け解析によって解析した結果において、文中の事象表現に対して、事実性のラベル(表2)のいずれかを付与する課題である。事象表現の判定基準は、松吉ら[5]に従う。この判定を自動的に行うことは事実性解析の本質ではないため、形態素が事象表現であるかの判定は全て人手で行う。

3.2 事実性解析に関わる言語要素

先行研究によって、事実性解析に関連する言語要素は文内では大きく4つに分けられることが分かっている。事象表現に後続する機能表現、疑問詞を含む文中の副詞、文節境界を越えて事実性に影響を与える語とそのスコープ、その他(疑

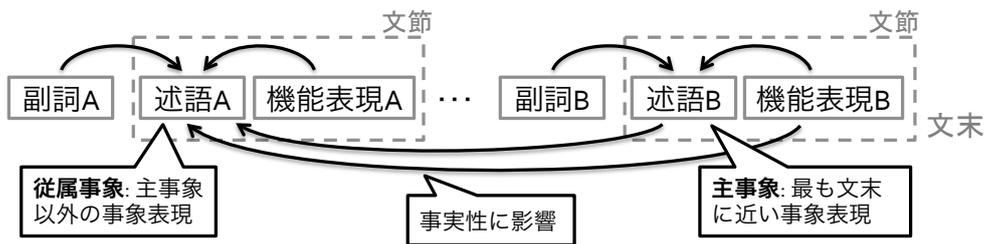


図 1: 事実性に関わる言語要素の構造: 矢印は、要素が事象表現の事実性に影響することを示す。

問符など)の4種類である。図1に、その他を除く3つの要素について、文内における関係を示す。図中の矢印は、その言語要素が事象表現の事実性に影響することを示している。事象表現は、文中での出現位置によって二種類に分類する。各文につき、最も文末に近い事象表現を主事象と呼び、それ以外の事象表現を従属事象と呼ぶ。関わってくる言語要素の種類が、主事象と従属事象とで異なることから、本研究では、主事象と従属事象とを分けて課題分析を行う。以下では、その他以外の3つの要素について、事象表現の事実性にどのように関連するかを述べる。なお、文をまたいだ言語要素による否定や推量も存在するが、本研究では文内の現象のみを取り扱う。

3.2.1 機能表現

事象表現に直接後続する機能表現(図1では、述語Aに対する機能表現A、および、述語Bに対する機能表現B)の問題は、多義性と多様性の二つに大きく分けられる。

- (9) a. 太郎は<<走っ>>たんでした よね_{態度}
- b. 太郎は<<走る>>んです よね_{疑問}

まず、多義性について、(9)に示す2つの例には、いずれも機能表現「よね」が出現しているが、示す意味は異なる。機能表現の意味は、下線部に続く下付文字で表す。(9a)は「太郎が走る」ことを推量しているが、(9b)は「太郎が走る」こ

とを確認していることから事象は成立していないことを示している．このような文を解析するためには，機能表現の多義性の解消は必須の技術である．

- (10) a. 太郎は << 走ら >> ない 否定。
b. 太郎は << 走る >> わけない 否定。
c. 太郎が << 走ら >> ねえ 否定。
d. 太郎が << 走れる >> もんか 否定。

次に，表現の多様性について，(10) に示す 4 つの例は，いずれも「太郎が走る」という事象が成立していないということを，異なる機能表現によって記述している．そのため，否定を認識するためには，典型的な否定の機能語である「ぬ，ない」だけでなく，「ねえ，もんか」といった砕けた表現もとらえる必要がある．

これらに加えて，複数の表現の組み合わせの問題がある．

- (11) a. 太郎は << 走れ >> なくなる 否定 ようだ 推量。
b. 太郎が << 走る >> かもしれない 推量。

機能表現の組み合わせは，(11a) のように複数の表現の意味が組み合わさって事実性を表す場合と，(11b) のように複数の形態素が組み合わさってはじめて意味を持つ場合（複合辞と呼ばれる）がある．(11a) は，否定の機能表現「なくなる」と，推量の機能表現「よう」が組み合わさることで「太郎が走る」という事象が成立しないことを推測していることを示している．この事例を正しく認識するためには，機能語単位での意味ラベルだけでなく，その組み合わせに従って事実性を演算することが必要となる．一方，(11b) は複合辞の事例であるが「走る」に後続する 3 つの単語「かも，しれ，ない」は，ひとまとまりで推量の機能表現を構成している．このとき「ない」は否定の意味を持っておらず，機能表現を解釈するには，特定の単語列を複合辞としてまとめた上でその意味を認識する必要がある．

3.2.2 述語周辺の副詞

事実性は、事象に後続する機能表現だけでなく、周辺の副詞（図1では、述語 A に対する副詞 A、および、述語 B に対する副詞 B）によって決定される場合がある。

- (12) a. 確か 太郎は << 走っ >> た。
b. 太郎は 果たして << 走る >> のだろうか。
c. どうしたら 太郎は << 走る >> だろう。

(12) に示す例は、いずれも下線部の副詞が事象「太郎が走る」の事実性に影響する。(12a) では、副詞がなければ事象は成立しているが、副詞「確か」が付加されることによって確信度が下がる。(12b) は、同様に「果たして」が付加されることにより、事象成立の確信度は大きく下がり、どちらかといえば事象は成立しないと読み取れる。(12c) は、下線部の副詞がなければ、推量を意味する機能表現「だろう」により、事象の成立を推量していると読み取れる。しかし方法を問う副詞「どうしたら」が付加されることにより、事象は成立していないと読み取れる。また、このとき「だろう」は推量の意味を持たず、疑問の意味となる。副詞は、用法がまとめられた辞書はあるものの [36]、事実性に及ぼす影響についての研究は進められていない。よって、副詞が事実性に影響を与える事例を収集するところから着手する必要がある。

3.2.3 文節境界を越えて事実性に影響を与える語とそのスコープ

事象表現が含まれる文節よりも文末側に現れる語（図1では、述語 A に対する述語 B および機能表現 B）によって、事実性が決定される場合がある。

- (13) a. 太郎は / << 走る >> / ことを / 拒否した。
b. 太郎は / << 走る >> と / 言っていたが、 / やめた。
c. 太郎は / << 走り >> も / 歩きも / しなかった_{否定}。

d. 太郎は / << 走った >> が、 / 楽しく なかった_{否定} らしい_{伝聞}。

(13) に「太郎が走る」という事象が、文節境界（” / ”で示す）を越えた後続の述語や機能表現によって、否定あるいは推量されている事例を示す。(13a) および (13b) は、下線部の述語によって事象の成立が否定されている。このような述語は、他の事象表現の事実性に及ぼす影響および範囲を決定することが重要である。(13c) は、後続の述語「歩く」に付随する否定の機能表現「なかった」が、事象「太郎が走る」にも影響して、その事実性が CT- であることが示唆される。一方で、(13d) は、後続の述語「楽しい」に否定の機能表現「なかった」と、伝聞の機能表現「らしい」が付随するが、事象「太郎が走る」の事実性には影響せず、この事象が成立することが読み取れる。このように、後続の述語に付随する機能表現が、文節境界を越えて事実性に影響する場合があります、その範囲の同定は、否定 / 推量のスコープの問題として知られている。

3.3 課題分析の方針

事実性は、3.2 節で述べた各要素が単体で影響するだけでなく、その組み合わせによって決定される。

(14) a. 太郎が << 走ら >> ない_{否定} というのは 間違っていた。(機能表現と後続する述語の組み合わせ)

b. たぶん 太郎は << 走ら >> ない_{否定}。(副詞と機能表現の組み合わせ)

例えば (14a) は、事象表現「走る」の直後にある否定の機能表現「ない」と、後続する述語「間違っていた」が組み合わせられて、事象「太郎が走る」が成立することを示している。(14b) は、副詞と機能表現の組み合わせによって、事象が成立しないことが推量されている。

課題分析においては、複合的に影響する要素は可能な限り切り分けることが重要である。3.2 節で述べた 3 種類の要素の中では、機能表現は、記述的研究に基づいて体系化が進められている領域であり [37, 38]、辞書も整備されている [29] ため、切り分けが容易であると考えた。そこで、事実性が機能表現のみで決定可能

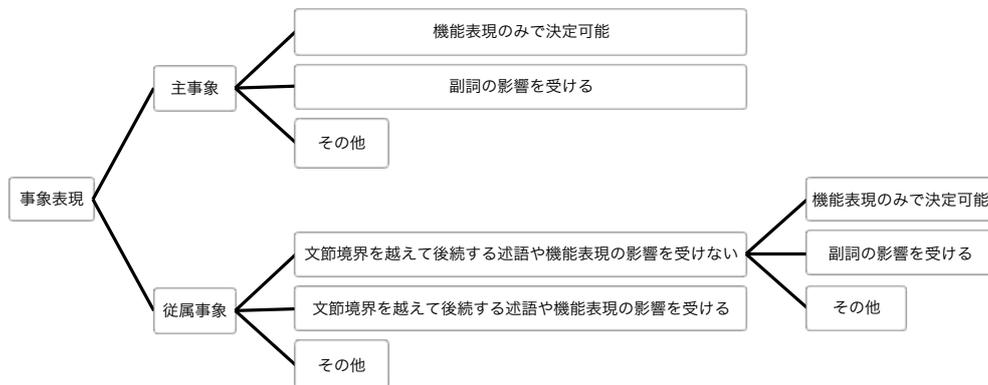


図 2: 事実性解析課題の切り分け

であるかという点と、文節境界を越えて後続する述語や機能表現の影響を受けるかという点の2つに着目して、課題を切り分ける。次に、主事象は、文節境界を越えて後続する述語や機能表現を持たないため、その影響を受けることはない。一方で従属事象は、後続の述語や機能表現の影響を受ける可能性がある。そこで、主事象と従属事象とに分けて誤り分析を行うことで、要素を切り分けた課題分析を行う。

機能表現の問題を切り分けるために、機能表現に基づくルールベースの事実性解析器を構築する（詳細は3.6節で述べる）。本解析器は、各事象表現について、直接後続する機能表現の意味ラベルのみに基づいて事実性を決定する。決定的に事実性を解析するモデルを適用し、その誤り事例のみを分析対象とすることで、難度の低い事例を分析対象から除外し、課題分析に注力することができる。

本解析器を主事象に対して適用すると、正解事例は機能表現のみで決定可能な事例であると判断できる。一方で、誤り事例は以下の3種類に分類できる。

1. 機能表現の意味ラベルや解析器のルールがナイーブであることが原因で誤ったが、機能表現のみで決定可能な事例
2. 副詞の影響を受けるため機能表現のみでは決定できない事例
3. その他

次に、従属事象に対して適用すると、正解事例は、主事象と同様に、機能表現の

みで決定可能な事例である。一方で、誤り事例は以下の4種類に分類できる。

1. 文節境界を越えて影響を及ぼす述語を持つ事例
2. 文節境界を越えて影響を及ぼす機能表現を持つ事例
3. 1. および 2. 以外で、機能表現の意味ラベルや解析器のルールがナイーブであることが原因で誤ったが、機能表現または副詞によって決定可能な事例
4. その他

まとめると、事実性解析課題の切り分けは図2のようになる。このような誤り分析によって、事実性解析の性能を向上させるには、どの言語要素に注力することが重要か、また各要素の部分課題にどの程度の解析性能が要求されるのかが明らかとなる。

3.4 事実性ラベル付与コーパス

事実性解析のエラー分析を行うためには、事実性ラベルが付与されたコーパスが必要である。本研究では、課題分析のために拡張モダリティタグ付与コーパス [5, 16] を利用する。拡張モダリティタグ付与コーパスは、『現代日本語書き言葉均衡コーパス』(BCCWJ)⁴ を付与対象としており、そのうちの Yahoo!知恵袋データを利用して誤り分析を行う。拡張モダリティタグ付与コーパスを用いるのは、事実性に関する情報が付与されており、課題分析に十分な事例数が確保できるとともに、同様の規模の利用可能なコーパスが他にないためである。拡張モダリティタグ付与コーパスには、Yahoo!知恵袋データの他にも、新聞、書籍、白書を対象としたデータも含まれているが、本研究では Yahoo!知恵袋データを対象に分析を行い、その他のドメインに対する分析は今後の課題とする。その理由としては以下の2点がある。

1. アノテーションガイドライン⁵の付録Bに「OC (Yahoo!知恵袋) の 14,089 事象に対しては、実装した解析システムの解析結果をフィードバックさせ、

⁴http://www.ninjal.ac.jp/corpus_center/bccwj/

⁵<http://cl.cs.yamanashi.ac.jp/nldata/modality/>

表 3: コーパス中の事実性の分布

| 事実性 | CT+ | PR+ | PR- | CT- | Uu | 合計 |
|------|-------|-----|-----|-----|-----|-------|
| 主事象 | 808 | 92 | 13 | 119 | 501 | 1,533 |
| 従属事象 | 1,538 | 177 | 36 | 163 | 287 | 2,201 |

それを参照しながらのタグ見直し作業を数回行い、タグの質を向上させている。」と記載がある。できるだけ信頼性の高いタグで誤り分析を行うため、Yahoo!知恵袋データを利用した

2. 言論マップ [39] や対災害 SNS 情報分析システム [40] などの、Web データを用いたアプリケーションに対する利用を想定している

拡張モダリティタグ付与コーパスには 6,362 文が収録されている。6,362 文のうち、主事象に機能表現を一つ以上含む文は 5,436 文ある。そのうちの約 30%である 1,533 文をランダムに選択し、課題分析の対象とした。この中には、主事象が 1,533 事象、従属事象が 2,201 事象含まれている。表 3 に、本実験の解析対象である 1,533 文における、主事象と従属事象の事実性の分布を示す。CT+がマジョリティではあるが、対象が Yahoo!知恵袋であるため、Uu も多く存在している。

3.5 機能表現意味ラベル付与コーパス

機能表現の多様性を吸収するために、拡張モダリティタグ付与コーパス中で、主事象に機能表現が付随する 5,436 文について、述部に続く機能表現に意味ラベルを付与したコーパスを構築する。3.4 節のコーパスと異なり、利用可能なコーパスおよびアノテーション仕様がないため、本研究で新たに構築した。事実性解析の課題分析の対象は、コアデータの約 30%であったが、機能表現解析の評価では、残る約 70%にも機能表現意味ラベルを付与した。構築したコーパスは、BCCWJ との差分データとして、アノテーション仕様と合わせて公開している⁶。まず 3.5.1 節で、機能表現の意味ラベルの設計方針について述べ、3.5.2 節では、設計した意

⁶<http://tinyurl.com/ja-fe-corpus>

表 4: 定義した機能表現意味ラベルの一覧

判断, 否定, 疑問, 推量-不確実, 推量-高確実性, 否定推量, 意志, 否定意志, 願望, 当為, 不可避, 依頼, 勧誘, 勧め, 許可, 不許可, 不必要, 無意味, 可能, 不可能, 困難, 容易, 自然発生, 自発, 無意志, 試行, 伝聞, 回想, 継続, 発継続, 着継続, 最中, 傾向, 事前, 完了, 事後, 結果状態, 習慣, 反復, 経歴, 方向, 授与, 受益, 受身, 使役, 付帯-続行, 付帯-並行, 同時性, 継起, 終点, 相関, 対比, 添加, 比較, 場合, 順接仮定, 順接確定, 逆接仮定, 逆接確定, 理由, 目的, 並立, 例示, 様態, 比況, 程度, 強調, 感嘆, 態度, 内容, 名詞化, 話題

味ラベル体系に基づいて実際に構築したコーパスについて述べる。そして 3.5.3 節では、自動解析可能な程度のラベル体系になっていることを確認するため、機械学習に基づく機能表現解析器を構築し、性能を評価する。

3.5.1 機能表現意味体系の設計

機能表現の意味ラベルは、記述的研究 [37, 38] に基づいて体系化された『つつじ』 [29] がある。しかし、今村ら [31] によると、『つつじ』に掲載されていない機能表現および意味ラベルが存在する。本研究では、『つつじ』に不足する機能表現や意味ラベルを拡充した。意味ラベルの付与対象は、事象表現、すなわち述語に付随する機能表現である。「太郎が」「会社に」のように名詞に付随する機能表現も存在するが、それらは付与対象としない。『つつじ』では 89 種類の意味ラベルが定義されているが、最終的に本研究では、72 種類の意味ラベルを定義した。『つつじ』に比べてラベルの種類数が減るのは、前述の通り、名詞に付随する機能表現の意味ラベルを定義しなかったためである。本研究で定義した意味ラベルの一覧を表 4 に示す。

3.5.2 機能表現意味ラベル付与コーパスの構築

拡張モダリティタグ付与コーパス中で、主事象に機能表現が付随する 5,436 文について、述部が続く機能表現に意味ラベルを付与したコーパスを構築した。なお、品詞体系は IPA 品詞体系に従う。テキストには、UniDic 品詞体系に基づく

形態素情報が付与されているが、これを構文解析器 CaboCha [41] を用いて再解析した結果を用いる。

意味ラベルは、複合辞内の位置を表す要素と意味を表す要素の組み合わせによって表現し、形態素単位で付与した。複合辞内の位置を表す要素として、以下で示す IOB2 フォーマット [42] を採用した。

I 機能表現を構成する先頭以外の形態素

O 機能表現に含まれない形態素

B 機能表現の先頭の形態素

機能表現の意味は、3.5.1 節で述べた 72 種類の意味ラベルを用いる。表 5 に意味ラベルの付与例を示す。「パソコン」「壊れ」などの内容語「。」などの記号には「O」を付与する。「た」のように一形態素のみの機能表現には「B:完了」のように付与し、「てしまっ」「かもしれない」のような複合辞の場合には、先頭の形態素に「B:推量-不確実」のように付与し、後続の形態素に「I:推量-不確実」のように付与する。なお、「が」は機能語に分類されるが、本研究で対象とする述部の機能表現ではないため「O」を付与する。

アノテーション作業は、形態素区切りの各文を提示するアノテーションシートを用いて行った。コーパス構築に関わるアノテーションは、言語学に詳しい一名の日本語母語話者が行い、ラベルの定義およびアノテーションガイドラインの作成は、付与作業を行いながら行った。

構築した機能表現意味ラベル付与コーパス中の文、形態素、述部、機能表現、複合辞の数を表 6 に示す。意味ラベルが付与された機能表現の延べ総数は 19,334 個、機能表現の異なり総数は 825 個であった。機能表現として意味ラベルが付与された形態素数は 27,190 個であった。すべての機能表現のうち、複数の形態素からなる複合辞は延べ 5,393 個であった。複合辞の占める割合は 27.9% (5,393/19,334) にのぼっており、機能表現を自動で認識する上で、複合辞の同定は大きな課題の 1 つであることを示している。

一致率をはかるために、ガイドライン完成後に、外部業者に委託して作業員 2 名によるアノテーションを行い、作業員間の意味ラベル付与の一致率を算出した。

表 5: 意味ラベルの付与例

| 形態素 | 意味ラベル |
|------|----------|
| パソコン | O |
| が | O |
| 壊れ | O |
| て | B:無意志 |
| しまっ | I:無意志 |
| た | B:完了 |
| かも | B:推量-不確実 |
| しれ | I:推量-不確実 |
| ない | I:推量-不確実 |
| 。 | O |

表 6: 機能表現意味ラベルコーパスの統計情報

| | |
|----------------|--------|
| 文数 | 5,436 |
| 形態素数 | 97,943 |
| 述部数 | 11,594 |
| 機能表現数 (number) | 19,334 |
| 機能表現数 (type) | 825 |
| 複合辞数 | 5,393 |

アノテーション対象はランダムに選択した 700 文である。ただし、付与作業の時間とコストを考慮し、アノテーション対象は主事象に付随する機能表現のみに限定した。

一致率は、機能表現抽出の一致、意味ラベルの一致の 2 つの観点から κ 値 [43] を算出する。機能表現抽出の一致については、意味ラベルを問わず、どの部分が機能表現かの同定における一致を評価する。意味ラベルの一致は、両名が機能表現であると判定した形態素列に対して付与された意味ラベルが一致するかを評価する。算出された κ 値を表 7 に示す。

表 7: 作業員間一致率の κ 値

| 項目 | Kappa |
|--------|--------|
| 機能表現抽出 | 0.9708 |
| 意味ラベル | 0.8514 |

機能表現抽出の一致率は非常に高く、機能表現の切り分けは難しくないことが分かった。不一致のほとんどは、「なぜ先生と呼ばれますか」のような文において、「ます」を複合辞とするか個別の機能表現とするかの揺れによるものであり、ガイドラインへの記述によって改善が見込める。

意味ラベルについても、十分な一致率がみられた。意味ラベルの一致についてより詳細に分析を行うために、一方の作業員の付与結果を正解、もう一方の付与結果をシステムの出力とみなした場合の F 値による、作業員間一致率の評価も行った。評価は機能表現を一単位として行う。すなわち複合辞は、構成する形態素全てが一致する場合のみ正解とする。評価結果を表 8 に示す。全体として 8 割以上の一致率を達成したが、一致率の低い意味ラベルも見られた。特に、〈方向〉〈名詞化〉〈受益〉〈内容〉〈比較〉〈不必要〉の 6 ラベルは全ての事例が不一致であった。

その他に一致率が低かったのは、〈結果状態〉〈習慣〉〈継続〉といった意味ラベルである。(15) に、これらのラベルが付与された例を示す。意味ラベルは、2 人のアノテータの付与結果を併記する。これらはいずれも、(15) のように「ている」という見出し語に付与される意味ラベルであり、表層が同じ表現における曖昧性による不一致である。改善のためには精緻なガイドライン設計が必要となるが、これらの意味は文脈に大きく依存し、分類方法を明確に記述することが難しく、今後の重要な課題である。

- (15) a. ... 今どこまで進行してます_{結果状態/継続} か？
 b. ...7 時間睡眠を繰り返しております_{結果状態/習慣}。
 c. ... どのくらいの期間劇場で公開してます_{習慣/継続} か？

表 8: 機能表現意味ラベルの作業者間一致率

| ラベル | 適合率 | | 再現率 | | F_1 (%) |
|--------|-------|---------------|-------|---------------|-----------|
| 疑問 | 92.81 | (297/320) | 94.59 | (297/314) | 93.69 |
| 判断 | 92.86 | (247/266) | 95.37 | (247/259) | 94.10 |
| 完了 | 80.85 | (114/141) | 93.44 | (114/122) | 86.69 |
| 結果状態 | 54.60 | (89/163) | 74.79 | (89/119) | 63.12 |
| 習慣 | 89.47 | (34/ 38) | 40.00 | (34/ 85) | 55.28 |
| 態度 | 90.79 | (69/ 76) | 88.46 | (69/ 78) | 89.61 |
| 否定 | 76.47 | (52/ 68) | 70.27 | (52/ 74) | 73.24 |
| 受身 | 100.0 | (39/ 39) | 92.86 | (39/ 42) | 96.30 |
| 継続 | 71.43 | (10/ 14) | 25.64 | (10/ 39) | 37.74 |
| 話題 | 100.0 | (38/ 38) | 97.44 | (38/ 39) | 98.70 |
| 無意志 | 82.93 | (34/ 41) | 100.0 | (34/ 34) | 90.67 |
| 勧め | 76.92 | (10/ 13) | 34.48 | (10/ 29) | 47.62 |
| 理由 | 100.0 | (21/ 21) | 91.30 | (21/ 23) | 95.45 |
| 願望 | 100.0 | (12/ 12) | 85.71 | (12/ 14) | 92.31 |
| 自然発生 | 21.15 | (11/ 52) | 78.57 | (11/ 14) | 33.33 |
| 依頼 | 73.33 | (11/ 15) | 100.0 | (11/ 11) | 84.62 |
| 方向 | 0.00 | (0/ 0) | 0.00 | (0/ 11) | 0.00 |
| 不確実 | 100.0 | (10/ 10) | 100.0 | (10/ 10) | 100.0 |
| 着継続 | 47.37 | (9/ 19) | 100.0 | (9/ 9) | 64.29 |
| 許可 | 100.0 | (5/ 5) | 83.33 | (5/ 6) | 90.91 |
| 試行 | 85.71 | (6/ 7) | 100.0 | (6/ 6) | 92.31 |
| 可能 | 62.50 | (5/ 8) | 100.0 | (5/ 5) | 76.92 |
| 伝聞 | 100.0 | (5/ 5) | 100.0 | (5/ 5) | 100.0 |
| 当為 | 50.00 | (3/ 6) | 75.00 | (3/ 4) | 60.00 |
| 様態 | 100.0 | (4/ 4) | 100.0 | (4/ 4) | 100.0 |
| 程度 | 60.00 | (3/ 5) | 100.0 | (3/ 3) | 75.00 |
| 強調 | 100.0 | (3/ 3) | 100.0 | (3/ 3) | 100.0 |
| 自発 | 40.00 | (2/ 5) | 100.0 | (2/ 2) | 57.14 |
| 名詞化 | 0.00 | (0/ 2) | 0.00 | (0/ 2) | 0.00 |
| 受益 | 0.00 | (0/ 0) | 0.00 | (0/ 2) | 0.00 |
| 意志 | 66.67 | (2/ 3) | 100.0 | (2/ 2) | 80.00 |
| 内容 | 0.00 | (0/ 1) | 0.00 | (0/ 2) | 0.00 |
| 目的 | 100.0 | (1/ 1) | 50.00 | (1/ 2) | 66.67 |
| 例示 | 100.0 | (1/ 1) | 100.0 | (1/ 1) | 100.0 |
| 比較 | 0.00 | (0/ 0) | 0.00 | (0/ 1) | 0.00 |
| 不必要 | 0.00 | (0/ 0) | 0.00 | (0/ 1) | 0.00 |
| 容易 | 100.0 | (1/ 1) | 100.0 | (1/ 1) | 100.0 |
| マイクロ平均 | 81.82 | (1,148/1,403) | 81.82 | (1,148/1,403) | 81.82 |

3.5.3 機能表現解析

前節までで機能表現に意味ラベルを付与したコーパスを構築したが，機械学習などで自動解析可能でなければ，本研究で提案する事実性解析は現実的でなくなってしまう．そこで，機械学習に基づく自動解析を行い，機能表現の意味ラベル付与課題について性能評価を行う．機能表現解析を系列ラベリング問題として定式化し，条件付き確率場 (Conditional Random Fields, CRF) [44] を用いる⁷．述語位置の判定は，事実性解析の設定と同様に，人手で与える．従って入力は，述語に続く機能表現列で，出力は意味ラベル系列である．コアデータ全体にあたる 5,436 文を用いて 10 分割交差検定によって性能を評価する．

学習素性には，主に，形態素解析で得られる単語素性を用いた．実験に使用した素性を「かもしれない」の「しれ」を例にして表 9 に示す．表 9 の $s, p, p_1, p_2, p_3, cf, bf$ はそれぞれ， n 番目の形態素 w_n の表層，品詞，品詞細分類 1，品詞細分類 2，品詞細分類 3，活用形，基本形を表し， s', p', cf', bf' は， $n - 1$ 番目の形態素 w_{n-1} における表層，品詞，活用形，基本形を表す．各素性とも，前後 2 単語までの周辺形態素からも抽出し，相対的な位置情報を用いて区別した．

機能表現解析器の評価結果を表 10 に示す．CRF モデルによる解析性能は，83.44% を達成しており，構築した機能表現意味ラベル付与コーパスは，機械学習によってある程度の分類性能を実現可能なものになっていることが確認できた．本節で構築した，機能表現意味ラベル付与コーパスを用いて，事実性解析の課題を，機能表現に関するものとそれ以外のものに切り分けていく．

3.6 誤り分析に用いる事実性解析モデル

課題分析のためには，解析器はその挙動が明確である必要がある．事実性解析の先行研究の多くでは，機械学習に基づく解析手法が提案されてきたが，機械学習ではその挙動が分かりづらい場合がある．そこで，機能表現に基づくルールベースの解析器を提案する．ここで構築する事実性解析器は，機能表現の意味的抽象化および事実性の構成性（相互作用モデル）に基づく解析手法であり，課題分析

⁷CRF の実装には CRFSuite [45] を用いた．

表 9: 機能表現解析における学習素性の一覧および素性抽出例

| | テンプレート | 説明 | 例 |
|---|--|----------------------------------|-----------|
| unigram 素性 | {s} | 表層形 | しれ |
| | {p} | 品詞 | 動詞 |
| | {cf} | 活用形 | 未然形 |
| | {bf} | 基本形 | しれる |
| | {p, p ₁ } | 品詞 + 品詞細分類 1 | 動詞_自立 |
| | {p, p ₁ , p ₂ } | 品詞 + 品詞細分類 1 + 品詞細分類 2 | 動詞_自立_* |
| | {p, p ₁ , p ₂ , p ₃ } | 品詞 + 品詞細分類 1 + 品詞細分類 2 + 品詞細分類 3 | 動詞_自立_** |
| | {bf, p} | 基本形 + 品詞 | しれる_動詞 |
| | {bf, p, p ₁ } | 基本形 + 品詞 + 品詞細分類 1 | しれる_動詞_自立 |
| {bf, p, p ₁ , p ₂ } | 基本形 + 品詞 + 品詞細分類 1 + 品詞細分類 2 | しれる_動詞_自立_* | |
| bigram 素性 | {w', w} | 基本形 bigram | かも_しれ |
| | {p', p} | 品詞 bigram | 助詞_動詞 |
| | {bf', bf} | 基本形 bigram | かも_しれる |
| | {cf', s} | 活用形 + 表層形 | *_しれ |

表 10: 機能表現解析器の評価結果

| 精度 | 再現率 | F ₁ |
|-----------------------|-----------------------|----------------|
| 83.20 (16,179/19,446) | 83.68 (16,179/19,334) | 83.44 |

がしやすい挙動が明確なモデルである。また、機能表現の意味を素性として学習を行った研究 [9, 16] は既に存在しているものの、それらをルールによって組み合わせ、決定的に事実性解析を行った手法は今までになく、新規性の高い解析手法となっている。事実性解析器の入力は、形態素解析された文、解析対象となる事象表現すなわち述語の位置、および後続する機能表現の意味ラベルである。機能表現の意味ラベルは人手で付与したラベルを与える。出力は、事象表現に対する事実性ラベルである。形態素情報は、UniDic 体系で与えられているが、機能表現は IPA 辞書体系であるため、オフセット情報に基づいて、UniDic 体系に自動的に対応づける。

本研究では、事実性の解析に、各事象表現よりも後ろにある機能表現の意味ラ

Algorithm 1 ルールベースの事実性解析モデル

```
1: Input: 解析対象となる事象の核となる形態素 (事象表現), 文全体の形態素情報, 事象表現に付随する機能表現の意味ラベル列
2: Output: 事実性ラベル
3: INITIALIZE  $C \leftarrow CT, P \leftarrow +$ 
4: for all 付随する機能表現の各意味ラベル (文末側から順に) do
5:   if 意味ラベルが更新ルール 1 に該当する意味ラベルのいずれか then
6:     if  $P = +$  then
7:        $P \leftarrow -$ 
8:     else if  $P = -$  then
9:        $P \leftarrow +$ 
10:    end if
11:  end if
12:  if 意味ラベルが更新ルール 2 に該当する意味ラベルのいずれか then
13:    if  $C = CT$  then
14:       $C \leftarrow PR$ 
15:    end if
16:  end if
17:  if 意味ラベルが更新ルール 3 に該当する意味ラベルのいずれか then
18:     $C \leftarrow U, P \leftarrow u$ 
19:  end if
20: end for
21: return  $C, P$ 
```

ベルを利用する。例えば、「〈否定〉の機能表現が付随する場合は肯否極性を反転」といった事実性更新ルールを適用する。主事象の事実性は、文末から主事象の間に存在する、すべての機能表現の意味ラベル列に基づいて更新ルールを適用することで決定される⁸。従属事象の事実性は、従属事象から次の内容語までの間に連なる機能表現の意味ラベル列に基づいて更新ルールを適用することで決定される。更新ルールは、機能表現意味ラベルの定義に基づいて、以下の3つを作成した。

1. 機能表現の意味ラベルが〈否定〉〈否定意志〉〈否定推量〉〈無意味〉〈不明確〉〈不可能〉〈回避〉〈不必要〉〈放置〉〈困難〉のいずれかの場合、肯否極性を+の場合は-に、-の場合は+に反転する

⁸疑問符などの記号も事実性に影響を与える要素として考えられるが、疑問符があっても、事実性がCT+である事象も少なくないため、本研究では採用していない。

2. 機能表現の意味ラベルが〈推量-不確実〉〈推量-高確実性〉〈否定推量〉〈意志〉〈否定意志〉〈伝聞〉〈様態〉〈容易〉〈困難〉のいずれかの場合，確信度を CT から PR に下げる
3. 機能表現の意味ラベルが〈疑問〉〈勧誘〉〈勧め〉〈願望〉〈依頼〉のいずれかの場合，事実性を Uu にする

それぞれの意味ラベルにおける表現例と，コーパス中にその意味ラベルをもつ機能表現が出現した延べ数を表 11 に示す．延べ数が 0 の意味ラベルは，分析対象のコーパスに一度も出現していない意味ラベルである．

ルールベースモデルのアルゴリズムを Algorithm 1 に示す．本モデルは，事象に付随する機能表現に基づく更新ルールを順次適用することで，事象の事実性を決定するモデルとなっている．以下にこのアルゴリズムによる解析例を示す．

- (16) 小さい方がいい場合も << ある >>らしい_{伝聞}ので_{理由} 一概にそうとも << 言え >>ない_{否定} みたい_{推量-不確実} です_{判断}。

主事象「言う」の事実性を決定する場合には，付随している 3 つの機能表現「ない」「みたい」「です」の意味ラベル列である〈否定〉〈推量-不確実〉〈判断〉に基づいて解析を行う．Algorithm 1 中の C, P は，それぞれ確信度，肯否極性の値をもつ変数であり，最終的にこれらの組み合わせで事実性の値を表す．初期値として， C に CT， P に + を割り当てる (line 3)．次に，文末側から順に，機能表現の意味ラベルに対応した更新ルールを適用していく (line 4-20)．まず「です」は〈判断〉の機能表現であり，更新ルール 1-3 のいずれにも該当しないため， C, P は更新しない．次に「みたい」は〈推量-不確実〉の機能表現であり，更新ルール 2 に該当するため， C を PR に更新し， P は更新しない (line 12-16)．最後に「ない」は〈否定〉の機能表現であり，更新ルール 1 に該当するため， C は更新せず， P を - に更新する (line 5-11)．結果的に， $C = PR, P = -$ となり，主事象「言う」の事実性として PR- が得られる (line 21)．従属事象「ある」の場合は，直後に連なる機能表現列である「らしい」「ので」の意味ラベル列〈伝聞〉〈理由〉に基づいて，更新ルール 2 のみを適用する (line 12-16)．その結果，従属事象「ある」の事実性は PR+ となる．ここで構築したモデルは，機能表現を意味ラベル

表 11: 更新ルールと意味ラベルの対応

| 更新ルール | 意味ラベル | 該当する機能表現例 | コーパス中の延べ数 |
|----------------|---------|-----------------|-----------|
| 1. 肯否極性を反転 | 否定 | ない, め | 323 |
| | 否定推量 | まい | 0 |
| | 無意味 | てもしかたがない | 0 |
| | 否定意志 | まいとする | 0 |
| | 不明確 | となく | 0 |
| | 不可能 | わけにいかない, ていられない | 5 |
| | 回避 | ところだった | 0 |
| | 不必要 | ことはない | 4 |
| | 放置 | ないでいる | 0 |
| | 困難 | にくい | 5 |
| 2. 確信度を減少 | 推量-不確実 | かもしれない | 34 |
| | 推量-高確実性 | だろう | 46 |
| | 否定推量 | まい | 0 |
| | 意志 | つもり, ようにする | 27 |
| | 否定意志 | まいとする | 0 |
| | 伝聞 | そう | 27 |
| | 様態 | っぽい | 11 |
| | 容易 | やすい | 11 |
| | 困難 | にくい | 5 |
| 3. 事実性を Uu に更新 | 疑問 | か, かな | 458 |
| | 願望 | てほしい, たい | 53 |
| | 勧誘 | うか, ないか | 3 |
| | 勧め | ては, ほうがよい | 68 |
| | 依頼 | ください, てもらえるか | 162 |

によって抽象化し, その意味ラベルに基づいて構成的に事実性を解析するモデルになっている。

ルールベースによる解析モデルを主事象に適用し, 誤り分析を行うことで, 機能表現のみで事実性が決定可能な事例の割合を明らかにするとともに, 副詞の影響を受ける事例がどの程度存在するのか, また, その他の要素はどのようなものがあるのかを明らかにする。次に, ルールベースによる解析モデルを従属事象に適用し, 誤り分析を行うことで, 機能表現以外的事实性を決定するための要素に関して, その重要性を定量的に分析し, 事実性解析の今後の方針を議論する。

また、本モデルは機能表現の意味ラベルのみを用いたシンプルなモデルであるため、必要以上に多く誤解析してしまう恐れがある。そこで、既存の素性 [9] を、オープンソースのモダリティ解析器 Zunda [46]⁹ に実装することで、リファレンスとなる解析性能を得る。Zunda は、拡張モダリティタグ体系に基づいて、タグごとに線形分類器による多クラス分類を行う。まず、真偽判断タグのラベルを表 2 に基づいて本研究の事実性ラベルに置き換える。他の 5 種類のタグについては、拡張モダリティタグをそのまま採用する。次に、素性は、江口ら [9] で利用されている素性のうち、リソースが利用可能なものを利用する。表 12 に、利用した素性の一覧と (16) から抽出される素性の例を示す。「事象選択述語が示唆する事実性」は、5.1 節で詳述するが、解析対象の素性として述語が含まれる文節の係り先文節に含まれる述語が示唆する事実性である。例えば「たばこを / << 吸う >> のを / << やめる >>。」について、「やめる」は係り元文節中の「吸う」が CT- であることを示唆する述語であることから、「吸う」を解析するとき、その事実性が CT- であることが示唆されるという素性を抽出する。最後に、分類器について、江口ら [9] は事象間の依存構造が考慮できる Factorial CRF [47] を利用していたが、Zunda は LIBLINEAR [48]¹⁰ を利用している。事象間の依存関係を考慮するため、解析対象の事象より文末側にあり、かつ最も近傍にある事象の拡張モダリティタグのうち、真偽判断と態度の 2 つについて、その解析結果を素性として利用する。例えば (16) では、解析対象が「ある」のときに、素性として「言う」の解析結果を利用する。LIBLINEAR の学習アルゴリズムは、L2 正則化ロジスティック回帰を利用し、パラメータは weight を 0 に設定した以外はデフォルトの値を利用した (epsilon= 0.1, cost= 1, bias= -1)。評価は 10 分割交差検定によって行う。文単位で分割することによって、同一文中の複数の事象が学習データとテストデータに属することはない。交差検定の段階では、主事象と従属事象は区別せずに学習させるが、精度と再現率を算出する段階では、主事象と従属事象を区別する。

⁹<https://code.google.com/p/zunda/>

¹⁰<http://www.csie.ntu.edu.tw/~cjlin/liblinear/> の 1.80 を利用した

表 12: 機械学習モデルで用いた素性一覧および (16) における素性抽出例

| | 素性 | (16) の主事象 「言う」に対する素性 | (16) の従属事象 「ある」に対する素性 |
|----------------------|---|--|---|
| 江口他 (2010) の素性 | 解析対象の事象を含む文節 (中心文節), その前後の文節, 中心文節に係る文節, 中心文節の係り先 2 つまでの文節 における形態素情報 (基本形, 品詞, 活用形) の 2-gram, 3-gram | 言え_ない, 言え_ない_みたい, そう_と, ある_らしい, 一概_に, そう_と 言え_ない ... | ある_らしい, ある_らしい_の, 場合_も, 一概_に 場合_も, ... |
| | 機能語列 (解析対象の事象以降で, 次の内容語が 出現するまでの形態素列) | らしい_ので | ない_みたいです |
| | 人手で付与した機能表現の意味ラベル | 伝聞, 理由 | 否定, 推量-不確実, 判断 |
| | 事象選択述語が示唆する事実性 | - | - |
| 依存関係に 関する素性 | 解析対象の事象より文末側で最も近い事象の <真偽判断> と <態度> の解析結果 | - | 「言う」の解析結果 (PR-, 叙述) |

4 事実性解析実験に基づく課題分析

本章では，3章で構築した事実性解析器の誤り分析を行うことで，事実性解析の課題分析を行う．前述の通り，主事象と従属事象では，事実性を決定するために考慮すべき要素が異なるため，主事象の分析結果を4.1節で，従属事象の分析結果を4.2節で述べる．

4.1 主事象に対する事実性解析

3.6節で構築した事実性解析器を主事象に対して適用し，誤り分析を行うことで，機能表現のみで決定可能な事象，副詞の影響を受ける事象，その他の3種類に分類する．対象となる事象は1,533事象あり，その解析結果を表13,14に示す．表13には，確信度と肯否極性を組み合わせた事実性の各ラベルにおける精度，再現率，F値，およびそれらのマイクロ平均，マクロ平均を示した．表14には，確信度と肯否極性の二軸それぞれにおける精度，再現率，F値，およびそれらのマクロ平均を示した．これらの結果から，機能表現のみを利用したシンプルなルールベースモデルであっても肯否極性は高い精度，再現率で判定可能であることが分かる．一方で，確信度については，PRの分類性能は高くない．また，機械学習ベースのモデルでは，機能表現などの素性も導入されているものの，十分な性能があげられていない．これは，事例数の偏りや，機能表現の多様性などの要因により，事実性解析が簡単な課題ではないことを示している．ルールベースモデルと機械学習ベースのモデルとを比較すると，全体の事例数が少なく，大きな偏りもあるため，機械学習ベースのモデルの方が若干不利ではあることを考慮しても，ルールベースモデルは，機械学習ベースのモデルと遜色ない性能を示している．このことから，本ルールベースのモデルの性能は極端に低いわけではなく，このモデルを用いた誤りを分析することで，事実性解析の課題分析を行うのは妥当であるといえる．

前述の実験で正解した事象は，機能表現のみで決定可能な事象であると判断することができる．残る240個の誤り事例を分析することによって，機能表現の意味ラベルあるいは事実性解析モデルが原因による誤り事例，副詞の影響を受ける

表 13: 主事象に対する事実性解析の評価

| システム | 事例数 | CT+ | PR+ | PR- | CT- | Uu | マイクロ平均 | マクロ平均 |
|--------|-------|-------------------|-----------------|----------------|-------------------|-------------------|-----------------------|-------|
| | | 808 | 92 | 13 | 119 | 501 | 1,533 | - |
| ルールベース | 精度 | 90.0 (711/790) | 50.6 (40/79) | 42.9 (3/7) | 79.5 (101/127) | 82.6 (438/530) | 84.3 (1,293/1,533) | 69.1 |
| | 再現率 | 88.0 (711/808) | 43.5 (40/92) | 23.1 (3/13) | 84.9 (101/119) | 87.4 (438/501) | 84.3 (1,293/1,533) | 65.4 |
| | F_1 | 89.0 | 46.8 | 30.0 | 82.1 | 85.0 | 84.3 | 66.6 |
| 機械学習 | 精度 | 79.6 (770/967) | 63.3 (19/30) | - (0/0) | 76.5 (88/115) | 90.0 (379/421) | 81.9 (1,256/1,533) | 61.9 |
| | 再現率 | 95.3 (770/808) | 20.7 (19/92) | 0.0 (0/13) | 73.9 (88/119) | 75.6 (379/501) | 81.9 (1,256/1,533) | 53.1 |
| | F_1 | 86.8 | 31.1 | - | 75.2 | 82.2 | 81.9 | 55.1 |

表 14: 主事象に対する事実性解析の各軸ごとの評価

| システム | 事例数 | 確信度 | | | | 肯否極性 | | | |
|--------|-------|---------------------|------------------|-------------------|-------|-------------------|-------------------|-------------------|-------|
| | | CT | PR | U | マクロ平均 | + | - | u | マクロ平均 |
| | 927 | 927 | 105 | 501 | | 900 | 132 | 501 | |
| ルールベース | 精度 | 91.3 (837/917) | 51.2 (44/86) | 82.6 (438/530) | 75.0 | 91.9 (799/869) | 83.6 (112/134) | 82.6 (438/530) | 86.1 |
| | 再現率 | 90.3 (837/927) | 41.9 (44/105) | 87.4 (438/501) | 73.2 | 88.8 (799/900) | 84.9 (112/132) | 87.4 (438/501) | 87.0 |
| | F_1 | 90.8 | 46.1 | 85.0 | 73.9 | 90.3 | 84.2 | 85.0 | 86.5 |
| 機械学習 | 精度 | 82.5 (893/1,082) | 63.3 (19/30) | 90.0 (379/421) | 78.6 | 85.6 (853/997) | 81.7 (94/115) | 90.0 (379/421) | 85.8 |
| | 再現率 | 96.3 (893/927) | 18.1 (19/105) | 75.6 (379/501) | 63.4 | 94.8 (853/900) | 71.2 (94/132) | 75.6 (379/501) | 80.5 |
| | F_1 | 88.9 | 28.1 | 82.2 | 66.4 | 89.9 | 76.1 | 82.2 | 82.8 |

事例，その他の事例に分類する．誤り分析の結果を表 15 に示す．

機能表現のみで決定可能な事例が 4 割以上と，まだ多く残されている．本実験で用いたルールは人手で構築しているため，ルールの改善の余地が残されている．このようなルールが改善でき，機能表現をしっかりと捉えることが出来るようになると，すでに正解できている 1,293 事例と合わせて，機能表現のみで 90.5% (1,387/1,533) の正解率をあげられることがわかった．

- (17) a. ビタミンは、野菜や海草から << 補給する >> べき 当為 です 判断。
 (正解ラベルに基づく解析：CT+，正解：Uu)

表 15: 誤りの種類の分布 (カッコ内は, 事実性のアノテーション誤りを除いた部分での誤りの割合)

| | | 事例数 |
|-----------------|----------------|------------|
| 機能表現のみで 決定可能 | 機能表現の更新ルールの不足 | 28 (14.0%) |
| | 機能語相当表現の扱い | 26 (13.0%) |
| | 機能表現のアノテーション誤り | 40 (20.0%) |
| 副詞の影響を受ける | 副詞と機能表現の組み合わせ | 70 (35.0%) |
| | 機能表現よりも副詞が支配的 | 14 (7.0%) |
| その他 | 機能表現の省略 | 18 (9.0%) |
| | 形態素解析誤り | 4 (2.0%) |
| 事実性のアノテーション誤り | | 40 - |
| 合計 | | 240 |

- b. 入札前に << 確認す >> べき 当為 でし 判断 た 完了 ね 態度。
 (正解ラベルに基づく解析: CT+, 正解: CT-)
- c. 大至急オーストラリアへ書類を << 送ら >> なくてはなりません 当為。
 (正解ラベルに基づく解析: CT+, 正解: PR+)

(17) は, 機能表現だけで事実性を決定できるものの, 現在の更新ルールが不足しているために誤った事例である。このようなルール不足に起因する誤りとしては, <当為> や <不許可> のように, 更新ルールを割り当てるべき意味ラベルを追加することで改善が期待できる事例が見られた。(17a) の機能表現「べき」「です」はそれぞれ <当為> <判断> の意味をもっているが, これは, <当為> に関する更新ルールが不足していたことによる誤りである。(17a) は <当為> を更新ルール3の適用対象に加えれば解決する問題である。今回の分析対象のコーパス中に, <当為> が付随している主事象は, 8 事例見られた。そのうち, 更新ルール3を変更することによって, もともと正解できていた事例が 2 事例, 誤りだった事例が正解できるようになる事例が 4 事例, もともと誤っており, 更新ルール3を変更しても正解できない事例が 2 事例あり, 正解できていた事例が誤りとなるような事例は見られなかった。更新ルール3を変更しても正しく解析できない事例を (17b), (17c) に示す。(17b) では「確認する」に付随する機能表現列は, <当為> <判断> <完了> <態度> であるから, その事実性は CT+ となるが, 正解は CT- である。こ

の場合，更新ルール1の適用対象として〈当為〉と〈完了〉の組み合わせを追加し，更新ルール3は適用しないように変更することで正しく解析することができる．〈当為〉と〈完了〉の両方が付随する事例は，コーパス中に1事例のみであることから，この変更による悪影響はない．本分析では，1,533文を分析したが，このように一度しか現れない機能表現のパターンがある．従って，更新ルールを洗練するには，規則ベース，学習ベースのいずれのアプローチをとるにせよ，機能表現の意味ラベルのアノテーションを拡充していく必要があるであろう．(17c)は，「大至急」という副詞があることからPR+と判断されている．この事例は〈当為〉を考慮するだけでは不十分で，「大至急」という副詞を考慮しなければならない．一方で，副詞の影響を加味する必要がある事例は半分近くにはのぼった．

(18) a. やはり この御時世、〈〈きつい〉〉のではないでしょうか疑問？

(正解ラベルに基づく解析: Uu, 正解: PR+)

b. どうやって〈〈判別し〉〉てるんでしょうか疑問？

(正解ラベルに基づく解析: Uu, 正解: CT+)

(18)は，機能表現だけは不十分であり，副詞と機能表現とを組み合わせる必要がある事例である．いずれの事例も，〈疑問〉の機能表現が後続しているため，主事象の事実性はUuと解析された．しかしながら，(18a)は，問いかけではあるものの，推量の意味合いが強いため，正解はPR+となっている．(18b)は，前提として起こった事象である「判別する」の方法を問う文であるため，CT+が正解である．このような事象の事実性を決定するためには，〈疑問〉の機能表現を利用するだけでは不十分であり，「やはり」や「どうやって」のような副詞を手がかりとし，それらを組み合わせる必要がある．

(19) おそらくただの〈〈見栄っ張り〉〉です。

(正解ラベルに基づく解析: CT+, 正解: PR+)

(19)では，事実性に影響を与えるような機能表現は付随していないが，その代わりに副詞「おそらく」によって事実性が決定されている．どのような副詞が事実性に影響を与えるかを分類し，手がかりとして捉える必要がある．

また、機能表現や副詞のみでは決定できない、その他の誤りとして、機能表現の省略による誤り事例が見られた。

(20) とれないので <<注意>> !!

(正解ラベルに基づく解析：CT+，正解：Uu)

例えば(20)では、事象「注意」で文が終わっており、機能表現が存在していないが、依頼の意味をもつ文であることが解釈できる。しかしながら、機能表現のみに基づいた解析では、機能表現が存在していないために依頼の意味を捉えられず、正しく解析することができない。そこで、文末の感嘆符など、機能表現以外の要素を利用して解析を行わなければならない。また、「注意」で文が終わる場合には依頼文であることが多いと予測できるため、事象自身の情報を利用することで、「注意」で終わる場合には依頼であると判定する、といったことが考えられる。

形態素解析やアノテーションの誤りについて、正しい情報が与えられた場合についても検証した。まず、形態素解析および機能表現のアノテーション誤りが解消された場合、事実性も正しく解析可能であることが分かった。また、事実性のアノテーション誤りについては、システムが出力したラベルの方が正しいことが分かった。

以上より、主事象の事実性解析については、機能表現の意味ラベルが正しく解析できれば、現在の意味ラベルの体系と本研究で用いた単純な規則だけでも、90%に近い正解率が得られることがわかった。3節で述べたように、現在の機能表現の意味ラベルは、既存の記述的研究に基づいた体系になっているが、これが事実性解析に最適な体系になっているかを評価することは容易ではない。しかし、現在の体系でも90%に迫る正解率が得られる余地があることは、この体系に基づく機能表現の解析モデルを研究開発することに一定の支持を与えるものと考えられる。今後は、[32]のような機能表現解析の研究に注力したい。

もう一つの大きな課題は副詞の扱いである。今回得られた誤りの半数近くは副詞に起因するものであった。意味解析における副詞の扱いは先行研究も乏しく、辞書の整備を初め、やるべき課題は多い。まずは事実性解析という切り口で、それに関連する情報に焦点を当ててリソースを設計・開発していく予定である。

表 16: 従属事象に対する事実性解析の評価

| システム | 事例数 | CT+ | PR+ | PR- | CT- | Uu | マイクロ平均 | マクロ平均 |
|--------|-------|-----------------------|------------------|---------------|------------------|------------------|-----------------------|-------|
| | | 1,538 | 177 | 36 | 163 | 287 | 2,201 | - |
| ルールベース | 精度 | 78.1 (1,491/1,910) | 73.8 (31/42) | 40.0 (2/5) | 80.5 (91/113) | 58.0 (76/131) | 76.8 (1,691/2,201) | 66.1 |
| | 再現率 | 96.9 (1,491/1,538) | 17.5 (31/177) | 5.6 (2/36) | 55.8 (91/163) | 26.5 (76/287) | 76.8 (1,691/2,201) | 40.5 |
| | F_1 | 86.5 | 28.3 | 9.8 | 65.9 | 36.4 | 76.8 | 45.4 |
| 機械学習 | 精度 | 77.4 (1,480/1,913) | 55.6 (25/45) | - (0/2) | 66.9 (89/133) | 53.7 (58/108) | 75.1 (1,652/2,201) | 50.7 |
| | 再現率 | 96.2 (1,480/1,538) | 14.1 (25/177) | 0.0 (0/36) | 54.6 (89/163) | 20.2 (58/287) | 75.1 (1,652/2,201) | 37.0 |
| | F_1 | 85.8 | 22.5 | - | 60.1 | 29.4 | 75.1 | 39.6 |

表 17: 従属事象に対する事実性解析の各軸ごとの評価

| システム | 事例数 | 確信度 | | | | 肯否極性 | | | |
|--------|-------|-----------------------|------------------|------------------|-------|-----------------------|-------------------|------------------|-------|
| | | CT | PR | U | マクロ平均 | + | - | u | マクロ平均 |
| | 1,701 | 213 | 287 | | | 1,715 | 199 | 287 | |
| ルールベース | 精度 | 81.8 (1,655/2,023) | 72.3 (34/47) | 58.0 (76/131) | 70.7 | 85.0 (1,659/1,952) | 84.8 (100/118) | 58.0 (76/131) | 75.9 |
| | 再現率 | 97.3 (1,655/1,701) | 16.0 (34/213) | 26.5 (76/287) | 46.6 | 96.7 (1,659/1,715) | 50.3 (100/199) | 26.5 (76/287) | 57.8 |
| | F_1 | 88.9 | 26.2 | 36.4 | 50.5 | 90.5 | 63.1 | 36.4 | 63.3 |
| 機械学習 | 精度 | 80.5 (1,648/2,046) | 57.4 (27/47) | 53.7 (58/108) | 63.9 | 84.6 (1,656/1,958) | 74.1 (100/135) | 53.7 (58/108) | 70.8 |
| | 再現率 | 96.9 (1,648/1,701) | 12.7 (27/213) | 20.2 (58/287) | 43.3 | 96.6 (1,656/1,715) | 50.3 (100/199) | 20.2 (58/287) | 55.7 |
| | F_1 | 88.0 | 20.8 | 29.4 | 46.0 | 90.2 | 59.9 | 29.4 | 59.8 |

4.2 従属事象における事実性解析

表 16, 17 に, 2,201 の従属事象に対して事実性解析器を適用した結果を示す。主事象の場合と比較すると, 全体の性能は下がっており, 従属事象の方が解析が難しいことがわかる。機械学習ベースのモデルと比較すると, 主事象の場合と同様に, ルールベースモデルが, 機能表現等が素性に入った機械学習ベースのモデルと遜色ない性能を示している。従属事象の場合においても, 本ルールベースのモデルの性能は極端に低いわけではなく, このモデルを用いた誤りを分析することで, 事実性解析の課題分析を行うのは妥当であるといえる。

表 18: 誤りの種類の分布（カッコ内は，事実性のアノテーション誤りを除いた部分での誤りの割合）

| 機能表現の付随 述語の種類 | | 有 | | 無 | | 計 |
|-----------------------|-------------------|----|------|----|------|------------|
| | | 名詞 | 名詞以外 | 名詞 | 名詞以外 | |
| 機能表現のみで 決定可能 | 機能表現の更新ルールの不足 | 1 | 10 | - | - | 11 (6.2%) |
| | 機能表現のアノテーション誤り | 2 | 13 | - | - | 15 (8.4%) |
| 副詞の影響を受ける | 文脈と機能表現の組み合わせ | 1 | 11 | - | - | 12 (6.7%) |
| | 「たぶん」などの副詞，助詞など | 1 | 7 | 2 | 0 | 10 (5.6%) |
| 文節境界を越えて 表現の影響を受ける | 後続する述語の影響 | 2 | 11 | 10 | 2 | 25 (14.0%) |
| | 後方の機能表現が影響する範囲 | 6 | 32 | 31 | 11 | 80 (44.9%) |
| その他 | 機能表現の省略 | 0 | 1 | 1 | 1 | 3 (1.7%) |
| | 対義語を同時に用いることによる否定 | 1 | 3 | 0 | 0 | 4 (2.2%) |
| | 事象参照表現が命令形 | 0 | 2 | 0 | 0 | 2 (1.1%) |
| | その他 | 2 | 8 | 3 | 3 | 16 (9.0%) |
| 事実性のアノテーション誤り | | 2 | 14 | 5 | 1 | 22 - |
| 合計 | | 18 | 112 | 52 | 18 | 200 |

従属事象において，機能表現以外に考慮すべき要素として，どのような要素が重要なのかを定量的に分析するために，正解ラベルを用いた場合の誤り分析を行う．主事象においては機能表現が付随している事象のみ扱ったが，機能表現が付随した主事象を含む文における従属事象を対象としているため，必ずしもすべての従属事象に機能表現が付随しているとは限らない．そこで，機能表現が付随している事象であるか否かをまず分類する．また，名詞述語なのか，動詞や形容詞といった述語なのかが，事実性解析の難易度に影響していると考えられるため，名詞述語なのか否かに従属事象を分類する¹¹．510の誤り事例の中から，200事例をランダムにサンプリングし，誤り分析を行った．

従属事象特有の誤りとしては，後続する述語の影響が事実性を決めている場合と，さらにその後ろの機能表現が事実性を決めている場合の2種類が考えられる．

(21) a. 安いものだと〈〈防水〉〉は怪しいです。

（正解：PR-，システム：CT+）

b. 物語を〈〈楽しみ〉〉つつ、〈〈冒険〉〉を堪能してください。

¹¹単純に事実性が付与された名詞をすべて名詞述語と分類するわけではないため，人手による分類を行っている．例えば「判別する」「判別できる」など「～する」「～できる」が付随する場合には，まとめて動詞述語として扱うが，「判別をする」といった場合には「判別」を名詞述語，「する」を動詞述語として扱い，「判別」が「する」の項になっていると判断する．

(「楽しむ」正解：Uu，システム：CT+，「冒険する」正解：Uu，システム：CT+)

(21) は、機能表現の正解ラベルを用いても正解できなかった従属事象の例である。まず (21a) の従属事象「防水する」は、事実性の正解が PR- であるが、「防水する」自身に付随している機能表現だけでこれが決まっているわけではなく、後続する述語であり、疑いをもっていることを示す表現である「怪しい」の影響が大きい。このように、従属事象に後続する述語が事実性を決めている場合を「後続する述語の影響」による誤りと分類した。このような事象の事実性を解くためには、どの述語が事実性にどういった影響を与えるかを分類する必要がある。次に (21b) の「楽しむ」および「冒険する」は、正解が Uu であるが、(21a) と同様に、それぞれに付随している機能表現だけでは事実性は決定できない。しかしながら、後続する述語である「堪能する」自身にそういった影響があるとはいえない。これは、「堪能する」を抽象的な述語「する」に置き換えた場合でも、「楽しむ」および「冒険する」の事実性が Uu と判断できることから明らかである。ではどういった要素が「楽しむ」および「冒険する」の事実性を Uu にしているかということ、「堪能する」に付随する機能表現「ください」が影響を与えているということが考えられる。このように、後続する述語ではなく、さらにその後ろに見られる機能表現が広く影響を与えているために、事実性をうまく解析できない事例を「後方の機能表現が影響する範囲」による誤りと分類した。このような事象の事実性を解くために、機能表現がどの事象までその影響を与えるのかを解析する必要がある。「後続する述語の影響」と「後方の機能表現が影響する範囲」とを区別する基準としては、後続する述語を「する」などの抽象的な述語に置き換えた場合に、事実性が変化するかどうかを考える。例えば (21a) において「怪しいです」を「しています」と置き換えた場合、事実性が全く異なってしまう。このような述語の置き換えを一つの判断基準として、誤りの分類を行った。

誤り分析の結果を表 18 に示す。主事象の場合と同様の誤りも見られたが、従属事象特有の誤りが誤り全体の 6 割を占めた。特に、機能表現が付随していない名詞述語においては、機能表現が影響する範囲を考慮すべき誤りが大半を占めていることから、機能表現が影響する範囲を捉えることの重要性を示している。以

降の節では、後続する述語や機能表現が影響する範囲のような従属事象特有の問題を解決するために、既存のコーパス中における現象を分析することで、今後の方針について議論する。

4.2.1 事象参照表現に後続する述語に関する分析

「あり得る」のような、事実性に影響を与える述語（以降、事象選択述語と呼ぶ）については、[9]が構築した辞書がある。[9]は、拡張モダリティを解析する手がかりとして利用するために、拡張モダリティに影響を与える表現を収録した辞書（以降、事象選択述語辞書と呼ぶ）を構築した。事象選択述語辞書は、行為・出来事を表す事象を必須格にとり得る述語を対象に、分類語彙表[49]に収録されている述語の中から、拡張モダリティに影響を与える8,580述語を収録している。事象選択述語辞書の項目の例を表19に示す。この辞書は、各述語が格にとる事象に与える影響を、直前の事象の時制および、述語の肯否極性ごとに収録している¹²。この辞書のうち、真偽判断の項目が、事実性解析に利用できると考えられる。

(22) a. 問題が〈〈発生する〉〉のを防いだ。

b. 問題が〈〈発生する〉〉のを防がなかった。

例えば(22)の「防ぐ」という述語は、(22a)のような肯定環境下では不成立、(22b)のような否定環境下では成立というように、事象「発生する」の肯否極性に影響を与える。(23)の「忘れる」は、直前の事象の時制を考慮した例である。

(23) 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。

事象「発言する」に対して、(23a)では過去に成立している事象であるが、(23b)では「発言する」ことが実際には起こっておらず、不成立である。

事象選択述語に関する問題は、このような既存の辞書を手がかりとして解決できると考えられる。現在の辞書のカバレッジを見積もるため、表18において、後

¹²この辞書は<http://bit.ly/ja-esp-dic>より入手可能である。

表 19: 事象選択述語辞書の記述例

| 見出し語 | 格 | 直前の事象の時制 | 述語自身の肯否極性 | 真偽判断 | 価値判断 | 評価極性 |
|------|---|----------|-----------|------------|------------|----------------|
| 防ぐ | を | 未来 | 肯定 否定 | 不成立 成立 | 働きかけ 許可 | ネガティブ ポジティブ |
| | | 非未来 | 肯定 否定 | - - | - - | - - |
| 忘れる | を | 未来 | 肯定 否定 | 不成立 成立 | - - | - - |
| | | 非未来 | 肯定 否定 | 成立 成立 | - - | - - |
| 知る | を | 未来 | 肯定 否定 | 高確率 高確率 | - - | - - |
| | | 非未来 | 肯定 否定 | 成立 成立 | - - | - - |

続する述語の影響が原因とされた誤りである 25 事例を対象に、事象選択述語辞書がカバーできているかどうか、を人手で分類した。例えば (21a) の「怪しい」といった述語が辞書中に登録されているかを判断する。そして、「怪しい」が辞書中に登録されている場合、登録されている情報を利用すれば正しく事実性ラベルを選択できるのか、即ち (21a) では、「直前の事象の時制が未来」であり、「述語自身の肯否極性が成立」である場合に、辞書に「真偽判断が低確率」と登録されているかどうか、を人手で判定した。このとき、直前の事象の時制や述語自身の肯否極性も人手で判定を行った。その結果、25 事例のうち 20 事例については、事象選択述語が辞書に収録されており、辞書の情報を利用すれば正しく事実性ラベルを選択できることがわかった。現在の辞書でも事実性解析の精度向上に貢献できることを示している。残りの 5 事例についても、現在の辞書には収録されていないものの、適切な情報が辞書に収録されていれば、辞書情報を用いて正しく事実性ラベルを選択することができる。現在の辞書でカバーできていた述語とカバーできていなかった述語を表 20 に示す。「気がある」「関係ある」などの複合表現が現在の辞書でカバーできていない傾向が見られ、こうした多様な表現の獲得が今後重要な課題として浮かび上がった。

表 20: 誤り事例における事象選択述語；カッコ内は25事例中の延べ数を示す

| 辞書に掲載されている述語 | 辞書に掲載されていない述語 |
|---|---|
| <p> 思う (3) 勧める (2) 忘れる 少ない ない 辞退する 疑う 怪しい おかしい 心配 予定 検討する 使う 期待する よい 言う 言い切る </p> | <p> あり得る 気がある 関係ある 暇がある 有無 </p> |

4.2.2 事象間の接続表現に基づくスコープに関する分析

(21b) のように、機能表現が直接付随する事象だけでなく、従属事象にまで影響を与えることによって解析に失敗した事例が、従属事象における誤りの4割以上を占めた。このような後方の機能表現が影響する範囲による誤りを解消するために、どのような情報が利用できるのかを分析する。本研究では、機能表現が影響する範囲を決定する問題を、機能表現のスコープを認識する問題として扱う。スコープとは「否定などの作用が及ぶ範囲」[13]であり、(24)では、角括弧で囲まれた範囲が否定を表す機能表現のスコープとなる。

(24) a. [仕事で<<行っ>>た] の ではない否定。

b. <<残念>>なことに、[鈴木さんは<<来>>] なかつ否定た。

(24a)では、「仕事で行った」という事象が否定されている。(24b)では、「鈴木さんは来た」という事象が否定されており、「残念である」という事象は否定されていない。一方で、(21b)では、主事象に付随する「ください」の影響が、従属事象である「楽しむ」「冒険する」にも影響を与える。現在のモデルでは、スコープを機能表現の直前の事象のみとして解析を行うため、(21b)は正解できなかった。そこで、スコープを必要に応じて広げ、機能表現の影響をスコープ内の事例に与えることで、後方の機能表現が影響する範囲による誤りを解消することができる。機能表現のスコープを広げるべき場合とそうでない場合とを認識するために、どのような情報が利用できるのか、それらの事例の割合はどの程度なのかを分析する。

[50]は、従属節内の要素の表れ方に基づき、従属節を接続助詞で分類している。

(25) a. [タバコを飲むが] ガンのことは心配していない。

b. [タバコを飲みながら] おしゃべりしている。

例えば(25a)では、従属節の述語的部分「飲むが」には、「飲まないが」「飲んだが」「飲みますが」「飲むだろうが」などのさまざまな要素に入れられる。一方(25b)では、「*飲まないながら」「*飲んだながら」「*飲みましながら」「*飲むだろうな

がら」などを用いることは出来ず、表れる要素が制限されている。これは、「～ながら」を伴う従属事象では、主事象に付随する機能表現が否定やモダリティなどを表しており、接続表現「～ながら」によってスコープが広がっていることを示唆している。また、[51] は日本語の時制節性に着目することで、[50] の分類がさらに分類できることを示している。[52] は、[50] の分類を一部修正し、その分類をもとに疑問の焦点やスコープに関して議論している。このように、主事象と従属事象をつなぐ接続表現の差によって、スコープの判断に接続表現を利用することが考えられる。そこで、実際にコーパス中に含まれる文を対象に、機能表現のスコープが従属事象にまで及んでいるかどうかを接続表現ごとに分類することで、スコープを考えるべき事例がどの程度存在するのか、接続表現がスコープ解析ならびに事実性解析に利用できるのか、を明らかにする。

我々が分析の対象とした従属事象の誤り 200 件のうち、後方の機能表現を考慮しなかったことによる誤りは 80 件あった（表 18 の「後方の機能表現が影響する範囲」）。これらは、(21b) の従属事象「楽しむ」の事実性のように、文節境界を越えた後方の機能表現（この例では「ください」）を事実性推定に考慮していないことによる誤りである。これらの従属事象の事実性は後方の機能表現の影響を受けるので、それぞれの従属事象は後方の機能表現のスコープの中に入っていることになる。上の 80 件の従属事象がそれぞれ後方の機能表現にどのように繋がっているかのパターンを調べると表 21 のような分布が得られた。主なパターンは次のとおりである。

直接の項 従属事象（「冒険」）が上位事象（「堪能し」）の項になっており、上位事象に付随する機能表現（「ください」）の影響を受けるパターン

(26) 物語を楽しみつつ、《《冒険》》を《《堪能し》》て ください 依頼。

テ形接続 従属事象（「活かし」）がテ形接続で後続事象（「働く」）に係っており、その後続事象の機能表現（「なかっ」）の影響を受けるパターン

(27) うまく《《活かし》》て《《働く》》ことができ なかっ 否定 た。

項を修飾 従属事象（「難しい」）が後続の事象表現（「ある」）の項になっている名詞（「試験」）を修飾しているパターン

(28) そんなに「難しい」試験が「ある」のでしょうか^{疑問}？

名詞述語を修飾 従属事象（「質問し」が名詞述語（「子かな」）の名詞を修飾しており，その名詞述語の機能表現（「かな」）の影響を受けるパターン

(29) 昨日楽譜何がいいって「質問し」た「子」かな^{疑問}。

これらのパターンについては，事実性解析時に後続の機能表現の影響を考慮する必要があるが，そのためには当該の従属事象が後続の機能表現のスコープ内にあるかどうかを正確に判別する必要がある．そこで，こうした機能表現のスコープの分布についてさらにデータを拡充して調査を行った．

拡張モダリティタグ付与コーパスのうち，2個以上事象が含まれており，かつ，主事象の事実性がCT+ではない文を140文ランダムに抽出した．主事象の事実性がCT+でない文では，主事象の後ろに何らかの機能表現が付随している場合が多いため，今回の分析目的にかなうと考えられる．140文中には事象表現が全部で440個含まれ，そのうち主事象が140個，従属事象が300個であった．この300個の従属事象を対象に，主事象に付随する機能表現のスコープ内に従属事象が入っているか，主事象と従属事象の間にどのような接続パターンが見られるかを人手で調査した．ただし，当該の従属事象が主事象から表層的に離れている場合は，隣接する場合にくらべて主事象に付随する機能表現のスコープ内には入りにくいと予測されるので，表22では，上記300個の従属事象をさらに主事象に隣接する事例140個とそれ以外の160個に場合分けして集計した．ここでいう「隣接」とは，係り受け関係にある事象の中で最も表層上近いものを指す．係り受けは，CaboCha [41] による自動解析結果を利用した．

まず，主事象から離れた従属事象160個について，従属事象が主事象と同じスコープ内に入っているかどうかを調べた．表22に示すように，スコープ内に入っている従属事象が11個，スコープ外にある従属事象が147個，後方の機能表現ではなく事象選択述語の影響を加味すべき事象が2個であり，「スコープ外」への偏りが極めて大きいことがわかった．すなわち，主事象から離れた従属事象が主事象の機能表現の影響を受けることは極めてまれで，その可能性を事実性解析プロセスの中で考慮しても精度のゲインはほとんど期待できない．

つぎに、当該従属事象が主事象に隣接している事例 140 個の分布を表 22 に示す。上段の「スコープ内が見られた表現」には、従属事象が主事象に付随する機能表現のスコープ内に入っている場合が一度でも観察された接続パターンを並べた。「～てから」のように表 21 に入っているが、上記 140 個の事例の中には出現しなかったものも含めてある。表 21 と表 23 を合わせると興味深い知見が得られる。表 21 の誤りを解消するためには、主として「直接の項」「テ形接続」「項を修飾」「名詞述語を修飾」などの接続パターンのスコープを決定する必要があるが、このうち「直接の項」をのぞく 3 つのパターンはいずれもスコープ内外の選択が高度に曖昧であり（例えば「テ形接続」は「スコープ内」が 5 件、「スコープ外」が 9 件）、これらのパターンのスコープを決定する課題に注力することに一定の効用があることがわかる。(30) に「テ形接続」でスコープ内外が異なる例を示す。

- (30) a. うまく <<活かし>>て <<働く>>ことができ なかつ否定た。(スコープ内)
- b. 諸事情が <<あつ>>て <<離婚する>>ことができ なかつ否定た。(スコープ外)

一方「直接の項」については、つねにスコープ内であると判断してもよい。また、「～が」「～ので」「～たら」などの接続パターンは「スコープ外」への偏りが大きく、決定的に「スコープ外」と決めても大きなリスクにはならない可能性がある。その他の接続パターンに関しても、ある程度の偏りが見られ、規則ベースで決めても問題はないと考えられる。離れた事象と比較して、隣接する事象のほうがスコープ内に入る場合が多いことから、事実性解析プロセスの中で隣接事象のスコープを考慮することによって、ある程度のゲインが期待できる。

隣接事象のスコープを考慮することが、事実性解析の性能改善に繋がるのかを検証するために、隣接事象のスコープを付与し、それを考慮した解析モデルを適用して、誤り分析を行う。まず、隣接事象対に対して同じスコープ内に入るかを人手で付与する。例えば (30a) では、「活かす」と「働く」は同じスコープ内に入ると付与し、(30b) では、「ある」と「離婚する」は同じスコープに入らないと付与する。次に、解析モデルを、スコープを考慮したものに拡張する。同じスコープに入ると付与された事象対について、前件の事象（文頭側の事象）については、自身に付随する機能表現の意味ラベル列に加えて、後件の事象（文末側の事象）

に付随する機能表現の意味ラベル列についても考慮して、事実性の更新ルールを適用する。例えば (30a) では「活かす」と「働く」が同じスコープ内であり、(30b) では「ある」と「離婚する」が同じスコープ内にはない、というアノテーションを行う。このアノテーションを利用し、3.4 節で述べた解析モデルを拡張することで、事実性の解析を行う。具体的には、「同じスコープ内である」と付与された事象対のうち、前件の事象については、前件の事象自身に付随する機能表現の意味ラベル列に加えて、後件の事象に付随する機能表現の意味ラベル列に基づいた更新ルールを適用することで、事実性を決定する。例えば (30a) では「活かす」と「働く」が同じスコープ内であるため、「活かす」の事実性を決定する際には、「活かす」自身の機能表現がもつ更新ルールを適用する（今回は更新ルールをもつ機能表現は付随していない）だけでなく、「働く」に付随する機能表現である「なかっ」がもつ更新ルール 1 も適用する。

1,533 文のうち、2 個以上事象が含まれており、かつ、主事象の事実性が CT+ ではない 441 文を抽出し、その中で係り受け関係にある 900 事象対に対してスコープのアノテーションを行った。その結果、同じスコープ内に入ると判断されたのは 120 事象対であった。これらの事象対のうち、後件の事象の事実性はスコープに関わらず変化しないが、前件の事象の事実性はスコープを利用することによって、後件の事象に付随する機能表現の影響を受けて変化する。前件の事象 120 事象における事実性解析の性能を表 24、事実性解析性能の変化を表 25 に示す。事例数の変化を見ると、改善事例が多く、36 事例見られたものの、スコープを考慮しても誤る事例も 51 事例見られた。しかしながら、その誤り原因を確認すると、51 事例のうち 32 事例は事実性のアノテーション誤りであり、システムは正しく事実性を解析することができていた。それ以外の事例において、スコープを考慮しても正解できなかつたものとしては、以下の事例がある。

(31) あなた自身が <<貯金する>> くせを <<つけ>> ないと 当為。

(「つける」正解：Uu，スコープ無：CT+，スコープ有：CT+)

(「貯金する」正解：Uu，スコープ無：CT+，スコープ有：CT+)

(31) では、「貯金する」と「つける」が同じスコープ内にあると判断された事例であるが、従属事象「貯金する」だけでなく、主事象「つける」も誤りとなって

いる。主事象「つける」が誤った原因は、主事象における誤り分析で述べたように、主事象に付随する機能表現である〈当為〉のルールが不足していることである。このルールが追加されれば、主事象の改善とともに、同一スコープ内の事象である「貯金する」も同時に正しく判定できるようになる。このように、主事象で見られた誤りを改善することで、スコープ内と判断された従属事象の性能改善にもつながる事例が19事例見られた。スコープを考慮した事実性解析を行うことで、CT+以外の性能、特に再現率を向上させることができるため、マクロ平均はスコープを考慮した方が大きく上回る性能となった。このことから、隣接事象対のスコープ判定を精緻に行うことが、事実性解析の性能向上に貢献することを確認することができた。

以上の観察を合わせると、次のことが言える。

- 調査した接続パターンのうち、誤りの半分近く (36/80) にあたる「直接の項」は我々の分析データを見る限り、全ての場合においてスコープ内に来るので、述語項構造解析の結果に基づいてスコープを広げることにより、事実性解析の性能を向上させることができる。
- 誤りのうち4割以上 (33/80) にあたる「テ形接続」「項を修飾」「名詞述語を修飾」等の接続パターンの場合には、スコープ内外の選択が高度に曖昧であり、これらのパターンのスコープを決定する課題に注力することに一定の効用があることがわかる。
- スコープを手で付与し、事実性解析に取り入れることで、CT+以外の性能、特に再現率を向上させることができるため、マクロ平均はスコープを考慮した方が大きく上回る性能を得られる。このことから、隣接事象対のスコープ判定を精緻に行うことが、事実性解析の性能向上に貢献することを確認できた。

表 21: 誤り事例における接続表現の分類

| 接続表現 | スコープに起因する誤り |
|---------|-------------|
| 直接の項 | 36 |
| テ形接続 | 20 |
| 項を修飾 | 8 |
| 名詞述語を修飾 | 5 |
| 連用中止 | 4 |
| ～と | 2 |
| ～てから | 2 |
| ～つつ | 1 |
| ～ながら | 1 |
| ～とは | 1 |
| 合計 | 80 |

表 23: 主事象と最も近い従属事象との間の接続表現の分類

| | 接続表現 | スコープ内 | スコープ外 |
|-----------------|---------|-------|-------|
| スコープ内が見られた表現 | 直接の項 | 15 | 0 |
| | テ形接続 | 5 | 9 |
| | 項を修飾 | 12 | 15 |
| | 名詞述語を修飾 | 4 | 4 |
| | 連用中止 | 1 | 0 |
| | ～と | 0 | 5 |
| | ～てから | 0 | 0 |
| | ～つつ | 0 | 0 |
| | ～ながら | 1 | 0 |
| | ～とは | 0 | 0 |
| | ～ば | 1 | 7 |
| スコープ内が見られなかった表現 | ～が | 0 | 24 |
| | ～ので | 0 | 9 |
| | ～たら | 0 | 6 |
| | ～けど | 0 | 3 |
| | ～ても | 0 | 2 |
| | ～後 | 0 | 1 |
| | ～上 | 0 | 1 |
| | ～次第 | 0 | 1 |
| | ～たり | 0 | 1 |
| | ～から | 0 | 1 |
| | ～とき | 0 | 1 |
| | ～ところ | 0 | 1 |
| | ～場合 | 0 | 1 |
| ～以外に | 0 | 1 | |
| 並列構造 | 0 | 2 | |
| 合計 | | 39 | 97 |

表 22: ランダムに抽出した 140 文中の従属事象の分布

| | 主事象に最も近い従属事象 | 主事象から離れた従属事象 | 計 |
|---------------|--------------|--------------|-----|
| スコープ内 | 39 | 11 | 50 |
| スコープ外 | 97 | 147 | 244 |
| 事象選択述語の影響を受ける | 4 | 2 | 6 |
| 計 | 140 | 160 | 300 |

表 24: スコープのアノテーションによる事実性解析性能

| | 事例数 | CT+ | PR+ | PR- | CT- | Uu | マイクロ平均 | マクロ平均 |
|---------|-------|-----------------|----------------|--------------|----------------|-----------------|------------------|-------|
| | | 51 | 10 | 1 | 11 | 47 | 120 | - |
| スコープなし | 精度 | 48.0 (47/98) | 100.0 (1/1) | - (0/0) | 50.0 (2/4) | 64.7 (11/17) | 50.8 (61/120) | 52.5 |
| | 再現率 | 92.2 (47/51) | 10.0 (1/10) | 0.0 (0/1) | 18.2 (2/11) | 23.4 (11/47) | 50.8 (61/120) | 28.7 |
| | F_1 | 63.1 | 18.2 | - | 26.7 | 34.4 | 50.8 | 28.5 |
| スコープを考慮 | 精度 | 70.4 (19/27) | 42.9 (3/7) | 0.0 (0/1) | 69.2 (9/13) | 52.8 (38/72) | 57.5 (69/120) | 47.0 |
| | 再現率 | 37.3 (19/51) | 30.0 (3/10) | 0.0 (0/1) | 81.8 (9/11) | 80.9 (38/47) | 57.5 (69/120) | 46.0 |
| | F_1 | 48.7 | 35.3 | - | 75.0 | 63.9 | 57.5 | 44.6 |

表 25: スコープのアノテーションによる事実性解析結果の変化

| スコープ利用前 \ スコープ利用後 | 正解 | 誤り |
|-------------------|----|----|
| 正解 | 33 | 28 |
| 誤り | 36 | 23 |

5 否定・推量・疑問のスコープとその自動解析

前章までで述べたとおり，事実性解析には，大きく4つの課題があり，「ない」「だろう」などの機能表現，「たぶん」「なぜ」などの副詞，「やめる」「始める」などの他の事象の事実性に影響を与える述語，他の事象の事実性に影響を与える機能表現の有無である．4つめの課題は，スコープ解析として知られている「ない」や「だろう」といった，否定や推量を表す機能表現は，まずそれが付随している述語に対して影響する．例えば，「太郎は〈〈走って〉〉会社に〈〈行く〉〉だろう推量」は，だろう推量 は，それが付随する述語「行く」が推量された事象であることを示す．それと同時に，「走る」についても，推量された事象であることを示す．このとき，機能表現「だろう」のスコープは，文頭までである．一方で，「太郎は〈〈走った〉〉が会社に〈〈遅刻する〉〉だろう推量」では，「だろう」が影響を与えるのは「遅刻する」のみであり，「走った」には影響しない．このとき，「だろう」のスコープは「遅刻する」までであり，「走った」はスコープの外であるという．

スコープについての言語学における厳密な定義は，我々の調べた範囲では明確ではない．日本語を対象としたスコープ解析は，川添ら [6] によるコーパス構築や，松吉 [35] による否定の焦点アノテーションにとどまり，大規模な事例分析や自動解析器の構築は行われていない．このような背景で，スコープ解析における課題は大きく2つある．

1. 日本語を対象としたスコープ解析の課題設計は自明ではない
2. スコープ解析に利用可能なデータが存在しない

これらの問題から，本章では，事実性解析への応用を念頭においた日本語スコープ解析課題を設計し，スコープを手で付与した，スコープ解析のためのコーパスを構築する．また，構築したコーパスを利用して，機械学習に基づくスコープ解析器を構築する．

まず課題設計について，英語を対象としたスコープ解析は，Shared Task [33, 34] などで取り組まれており，1. 否定や推量を示唆する表現 (cue) の発見，2. cue が影響する範囲 (scope) の同定，3. scope 内の事象の中心となる述語 (event) の発見，の3段階の課題として設計されている．一方で，日本語では同様のタスク

は実施されておらず、また利用可能なデータも存在しない。日本語でも同様の課題設計が可能であるが、本稿では 2. の scope の同定に注力するために、1. の cue は機能表現に限定した上で正解を与え、3. は実施しないものとする。特に、1. の cue の自動同定は、内容語にまで範囲を広げると、否定だけでも多様であることが分かっている [53]。従って、本稿におけるスコープ解析は、否定、推量あるいは疑問を示唆する機能表現について、それが付随する述語より文頭側に隣接する述語にも影響するか、すなわち機能表現のスコープ内であるかを二値分類する課題とする。文中に述語が 3 つ以上ある場合、最も文末側の述語に付随する機能表現が、他の 2 述語に影響を与える可能性もあるが、本研究で調査した範囲では、そのケースは極めて少ない。本稿は日本語を対象としたスコープ解析の第一歩であるため、まずは頻度が高くかつ基本的な現象である、文末に最も近い事象（主事象）とそれに隣接する事象（従属事象）について、従属事象が主事象中の機能表現のスコープ内にあるかを判定する課題に取り組む。

次に、スコープ解析に利用可能なデータについて、日本語均衡コーパスのウェブドメインである Yahoo!知恵袋を対象に、人手でスコープの付与を行った。事実性解析への応用を視野に入れているため、前章までで分析を行ってきたドメインと同じドメインのデータを用いた。本稿では、統計的機械学習に基づいてスコープ解析に取り組むが、その学習データとして、知恵袋のコアデータ以外から 3,589 文を無作為に抽出し、スコープを付与した。学習データには拡張モダリティタグおよび機能表現の意味ラベルは付与されていない。機能表現については、文献 [32] の自動解析器を用いる。自動解析器の性能は、精度 77%、再現率 79%と、十分な性能を持っている。

5.1 スコープの付与および事例分析

スコープ解析は、事実性解析を難しくしている問題の一つであるが、その理由は、事例の偏りにある。前章までで述べたとおり、日本語均衡コーパスのウェブドメインである Yahoo!知恵袋の 1,533 文について分析を行った結果、文中で隣接する 2 つの述語について、文末側の述語に付随する機能表現が他方の述語に影響を与えることは、全体の 13% (120/900) に過ぎない。従って、スコープは常に狭

い、すなわち機能表現はそれが付随する述語以外の述語に影響することはないと判断することで一定の解析性能を担保できるため、本課題に取り組む場合は高い精度を実現することが要求される。しかしながら、述語間の接続表現に基づいて事例を分類すると、て接続については、スコープを広く考えるべき場合が少なくない。

そこで、述語間が「て接続」である場合に限定してサンプリングし、述語間のスコープ内外をアノテーションする。例えば、以下の例の1はスコープ内であり、主事象〈〈働く〉〉に付随する否定の機能表現の影響が、従属事象〈〈活かす〉〉にも影響することで、〈〈活かす〉〉の事実性をCT-にしている。一方で、2はスコープ外であり、主事象に付随する二つの機能表現は、いずれも従属事象に影響を与えていない。

1. うまく〈〈活かし〉〉て〈〈働け〉〉なかった_{否定}
2. 諸事情が〈〈あっ〉〉て〈〈離婚〉〉でき_{可能}なかった_{否定}

主事象に推量の機能表現が続く事例を同様に示す。否定の場合と同様に1がスコープ内、2がスコープ外である。否定に比べると、従属事象も推量されているかを判定するのはやや難しく、後述の通り作業間の一一致率も否定に比べると低くなる。

1. 連絡がないと〈〈思っ〉〉て、〈〈待っ〉〉ている_{継続}かもしれません_{推量-不確実}
2. 意外性が〈〈あっ〉〉て、逆に〈〈良い〉〉かもしれません_{推量-不確実}

最後に、主事象に疑問の機能表現が続く事例を示す。事例を分析すると、否定や推量と比べて、疑問表現の挿入位置によってスコープ内外が変化する傾向にあるため、そのような情報を素性でとらえることが重要であると考えられる。

1. なぜ〈〈分け〉〉て〈〈書い〉〉ている_{結果状態}のですか_{疑問}？
2. 新しいHDDに〈〈交換し〉〉て、この後どう〈〈する〉〉のでしょうか_{疑問}？

以上の事例のように、主事象と間近の従属事象におけるスコープの内外の判定を、Yahoo!知恵袋のコアデータ以外の数千文に対してアノテーションをした。

表 26: 付与されたスコープの割合

| 機能表現 | スコープ内 | スコープ外 | 合計 |
|------|-------|-------|-------|
| 否定 | 127 | 2,051 | 2,178 |
| 推量 | 341 | 160 | 501 |
| 疑問 | 500 | 410 | 910 |



図 3: スコープ解析の入力例：係り受け解析結果に機能表現ラベルが付与されたもの

アノテーション結果を表 26 に示す。否定については、スコープ外にやや偏りがあるが、他の 2 つの機能表現については、偏りは大きくない。アノテーションの一致率について、表 26 の一部についてもう一名の作業者が付与し、その一致率を計ったところ、否定は 97.4% (150/154)、推量は 80.8% (118/146)、疑問は 84.3% (102/121) となり、十分な一致率があった。

5.2 機械学習に基づくスコープ自動解析

前節で構築したスコープ付与コーパスを用いて、機械学習による二値分類を行う。二値分類器は、ロジスティック回帰を用いた。実装としては、Classias [54] を用いた。機械学習に用いた素性を、表 27 に示す。ただし、BOS は文頭、EOS は文末、BOC は文節の始まり、EOC は文節の末尾を示し、PRED は述語を抽象化した記号として用いる。素性抽出例は、図 3 に示す例文に対して素性抽出をした例である。

表 27: スコープ解析の素性

| ID | 素性 | 素性抽出例 |
|----|---|---|
| a | 文全体の表層, 原形, 品詞, 表層 + 品詞 (1-3gram) | 足, と, シッポ, を, 洗っ, て, ..., BOS_足, 足_と, ... か_?, ?_EOS |
| b | 前件の事象を含む文節の表層, 原形, 品詞, 表層 + 品詞 (1-Ngram) | 洗っ, て, おい, て, BOC_洗っ, 洗っ_て, ..., て_EOC |
| c | 後件の事象を含む文節の表層, 原形, 品詞, 表層 + 品詞 (1-Ngram) | 乾かし, て, いる, ..., BOC_乾かし, 乾かし_て, ..., ?_EOC |
| d | 係り受け木上での, 事象間の表現の表層, 原形, 品詞, 表層 + 品詞 (1-Ngram) | て, おい, て, PRED_て, て_おい, て_PRED |
| e | 後件の事象から文末までの表現の表層, 原形, 品詞, 表層 + 品詞 (1-Ngram) | て, いる, の, です, ..., PRED_て, て_いる, ..., ?_EOS |
| f | dにおける, 機能表現単位での表層, 意味ラベル, 表層 + 意味ラベル (1-Ngram) | ておい, て, PRED_ておい, ておい_て, ..., 結果状態_順接確定, PRED_結果状態, 結果状態_順接確定, ..., ておい/結果状態, ... |
| g | eにおける, 機能表現単位での表層, 意味ラベル, 表層 + 意味ラベル (1-Ngram) | ている, のですか, PRED_ている, ている_のですか, ..., 習慣, 疑問, PRED_習慣, 習慣_疑問, ..., ている/習慣, ... |
| h | 前件の事象を含む文節と後件の事象を含む文節が隣接しているか否か (binary) | False |
| i | 副詞の位置 (表層上, 係り受け構造上のそれぞれで, 前件の事象の前に副詞が存在するか否か, 前件の事象と後件の事象の間に副詞が存在するか否か; それぞれ binary) | False . False , False , False |

5.3 評価実験

交差検定およびオープンテストによって, スコープ解析の性能評価を行う. オープンテストを行うため, 5.1 節においてスコープを付与したデータから, 評価用データをサンプリングし, 残りを訓練用データとする. 評価用データとしては, 各機能表現から 100 文ずつランダムに抽出した. 各データにおけるスコープの割合を表 28 に示す. まず, 訓練用データのみを用いて 10 分割交差検定を行うことで, 学習方法や, 各素性の効果の比較を行う. 次に, 訓練用データ全てで学習をし, 評価用データでテストを行ったオープンテストの結果を示す. 各評価結果は, 全事例中の正解できた事例の割合 (正解率) で示す.

5.1 節では, 否定, 推量, 疑問のそれぞれの機能表現が付随している事例に対

表 28: 訓練用データと評価用データにおけるスコープの割合

| | 機能表現 | スコープ内 | スコープ外 | 合計 |
|--------|------|-------|-------|-------|
| 訓練用データ | 否定 | 118 | 1,960 | 2,078 |
| | 推量 | 273 | 128 | 401 |
| | 疑問 | 440 | 370 | 810 |
| 評価用データ | 否定 | 9 | 91 | 100 |
| | 推量 | 68 | 32 | 100 |
| | 疑問 | 60 | 40 | 100 |

表 29: スコープの自動解析結果 (機能表現の意味ラベルごとに分割した場合とそうでない場合との比較); 太字は性能が高いもの

| | 否定 | 推量 | 疑問 | マイクロ平均 |
|----|-----------------------------|------------------------|------------------------|-----------------------------|
| 一括 | 0.9461 (1,966/2,078) | 0.726 (291/401) | 0.728 (590/810) | 0.8656 (2,847/3,289) |
| 分割 | 0.9466 (1,967/2,078) | 0.723 (290/401) | 0.738 (598/810) | 0.8680 (2,855/3,289) |

してスコープを付与した。これらの機能表現ごとに別々に学習を行うべきか、あるいは、機能表現を区別せず一括して学習を行うべきかは明らかでない。そこでまず、素性の検討を行う前に、すべてのデータを一括して学習する方法か、否定、推量、疑問という意味ラベルごとにデータを分割してそれぞれ学習する方法か、どちらの学習方法がスコープ解析課題において適切かを調査する。一括して学習する方法では、否定、推量、疑問の機能表現を区別せず、訓練用データ 3,289 文をすべて用いて 10 分割交差検定を行う。分割して学習する方法では、訓練用データにおける否定 2,078 文、推量 401 文、疑問 810 文に対して、それぞれ 10 分割交差検定を行う。評価は、一括して学習した場合でも、否定 2,078 文、推量 401 文、疑問 810 文のそれぞれの意味ラベルが付随する事例に分けた正解率と、全事例 3,289 文の合計の正解率によって行った。素性は表 27 のものをすべて用いた。表 29 に評価結果を示す。この結果から、疑問の場合はわずかに一括して学習した場合のほうが性能が高かったものの、全体として意味ラベルごとに分割してそれぞれ学習したほうが性能が良くなることを確認できた。以降では、意味ラベルごとに分割してそれぞれ学習を行った結果を示す。

表 30: スコープの自動解析結果 (アブレーションテスト); *は全素性利用時と比較して性能の低下が見られたもの

| 素性 | 否定 | 推量 | 疑問 | マイクロ平均 |
|--------|-----------------------|------------------|------------------|----------------------|
| 全素性 | 0.9466 (1,967/2,078) | 0.723 (290/401) | 0.738 (598/810) | 0.8680 (2,855/3,289) |
| - 素性 a | 0.9471 (1,968/2,078) | *0.706 (283/401) | *0.709 (574/810) | 0.8589 (2,825/3,289) |
| - 素性 b | *0.9461 (1,966/2,078) | 0.723 (290/401) | *0.733 (594/810) | 0.8665 (2,850/3,289) |
| - 素性 c | *0.9461 (1,966/2,078) | *0.721 (289/401) | 0.742 (601/810) | 0.8683 (2,856/3,289) |
| - 素性 d | 0.9466 (1,967/2,078) | 0.723 (290/401) | 0.738 (598/810) | 0.8680 (2,855/3,289) |
| - 素性 e | 0.9471 (1,968/2,078) | *0.721 (289/401) | 0.744 (603/810) | 0.8696 (2,860/3,289) |
| - 素性 f | 0.9471 (1,968/2,078) | *0.721 (289/401) | *0.736 (596/810) | 0.8674 (2,853/3,289) |
| - 素性 g | 0.9471 (1,968/2,078) | 0.733 (294/401) | 0.738 (598/810) | 0.8696 (2,860/3,289) |
| - 素性 h | *0.9461 (1,966/2,078) | *0.721 (289/401) | *0.727 (589/810) | 0.8647 (2,844/3,289) |
| - 素性 i | 0.9466 (1,967/2,078) | *0.718 (288/401) | 0.746 (604/810) | 0.8693 (2,859/3,289) |

表 31: スコープの自動解析結果 (素性ごとの比較)

| 素性 | 否定 | 推量 | 疑問 | マイクロ平均 |
|---------|----------------------|-----------------|-----------------|----------------------|
| 全素性 | 0.9466 (1,967/2,078) | 0.723 (290/401) | 0.738 (598/810) | 0.8680 (2,855/3,289) |
| 素性 a のみ | 0.9447 (1,963/2,078) | 0.706 (283/401) | 0.726 (588/810) | 0.8617 (2,834/3,289) |
| 素性 a, h | 0.9466 (1,967/2,078) | 0.708 (284/401) | 0.740 (599/810) | 0.8665 (2,850/3,289) |

次に、どの素性が効果的かを調べるため、アブレーションテストを行った。表 30 に結果を示す。否定の場合には、素性を取り除いても、ほとんど性能の変化は見られなかった。これは、スコープ内 118 事例、スコープ外 1,960 事例と、事例数に大きな偏りがあり、学習がうまく行えていないためだと考えられる。最も基本的な素性である素性 a のみで学習した結果を表 31 に示す。事例数に偏りがあるため、最も基本的な素性のみで学習を行っても、大きな差は見られなかった。推量の場合には、6 種類の素性で性能の低下が見られ、特に素性 a が有効な素性となっている。しかしながら、表 31 に示す通り、素性 a のみで学習を行っても、全素性に匹敵する性能は見られず、それぞれの素性が少しずつ効いていることがわかった。疑問の場合には、4 種類の素性で性能の低下が見られ、特に素性 a, h が有効な素性となっている。素性 a, h のみで学習した結果を表 31 に示す。疑問の場合には、素性 a と h のみでも全素性に匹敵する性能となっており、他の素性は冗長なものとなっていることを示している。

最後に、訓練用データ全てで学習をし、評価用データでテストを行ったオーブ

表 32: スコープの自動解析結果 (オープンテスト); ベースラインは素性 a のみを用いて学習したもの

| | 否定 | 推量 | 疑問 | マイクロ平均 |
|--------|----------------------|----------------------|----------------------|------------------------|
| 全素性 | 0.93 (93/100) | 0.76 (76/100) | 0.70 (70/100) | 0.797 (239/300) |
| ベースライン | 0.93 (93/100) | 0.68 (68/100) | 0.64 (64/100) | 0.750 (225/300) |
| 多数派に分類 | 0.91 (91/100) | 0.68 (68/100) | 0.60 (60/100) | 0.730 (219/300) |

ンテストの結果を表 32 に示す。比較対象として、機能表現などの情報を使わず、文全体の形態素情報といった最も基本的な素性のみを用いたモデル、即ち素性 a のみを用いて学習を行うモデルをベースラインとした。また、それぞれの場合において、すべての事例を多数派に分類した場合の性能も合わせて示す。即ち、否定の場合には全てスコープ外、推量、疑問の場合には全てスコープ内に分類した場合の性能である。否定の場合には、全素性を用いても、ベースラインと性能の変化が見られなかった。また、全て多数派であるスコープ内に分類した場合と比較しても、ほとんど性能の向上が見られていない。これは、スコープ内外の事例数が大きく偏っていることにより、素性を追加してもうまく学習を行えていないことを示している。推量、疑問の場合には、ベースラインや、多数派に分類した場合と比較して性能の向上が見られた。

全素性を用いてオープンテストを行った結果について議論する。表 33 に、オープンテストにおける Confusion Matrix を示す。否定については、スコープ内が正解であるものがスコープ外である、と出力された誤りのみであり、やはり事例数の偏りによって学習が難しくなっていると考えられる。ベースラインと全素性を利用した場合とで、分類結果が異なった例を (32) に示す。

- (32) a. 郵便の料金で、郵便局まで << 行っ >> て << 確かめる >> のですか_{疑問} ?
 (正解: スコープ内, 全素性: スコープ内, ベースライン: スコープ外)
- b. << 結婚し >> て何年に << なり >> ますか_{疑問} ?
 (正解: スコープ外, 全素性: スコープ内, ベースライン: スコープ外)

(32a) は、全素性を用いることで正解できるようになった事例である。この事例については、それぞれの事象を含む文節が隣接していることを反映した、素性 h

表 33: オープンテストにおける Confusion Matrix ; 太字は正解事例

| 否定 | | | |
|---------|-----------|-----------|-----|
| 出力 \ 正解 | スコープ内 | スコープ外 | 計 |
| スコープ内 | 2 | 0 | 2 |
| スコープ外 | 7 | 91 | 98 |
| 計 | 9 | 91 | 100 |
| 推量 | | | |
| 出力 \ 正解 | スコープ内 | スコープ外 | 計 |
| スコープ内 | 62 | 18 | 80 |
| スコープ外 | 6 | 14 | 20 |
| 計 | 68 | 32 | 100 |
| 疑問 | | | |
| 出力 \ 正解 | スコープ内 | スコープ外 | 計 |
| スコープ内 | 49 | 19 | 68 |
| スコープ外 | 11 | 21 | 32 |
| 計 | 60 | 40 | 100 |

が有効であった。一方，(32b) は，全素性を用いることで誤ってしまった事例である。この事例に関しては，どの素性が特に重いということではなく，様々な素性が誤りの原因となっていることがわかった。このことから，素性を再検討し，冗長な素性を削除する，といったことが必要となる。

5.4 スコープ解析のまとめ

事実性解析における課題を切り分けていく中で，特に従属事象においては，スコープ解析が最も大きな課題であることを明らかにしてきた。そこで本章では，スコープ解析における課題に取り組み，スコープ解析器を構築し，その評価実験を行った。

スコープ解析における課題は大きく 2 つある。

1. 日本語を対象としたスコープ解析の課題設計は自明ではない
2. スコープ解析に利用可能なデータが存在しない

1の課題設計に関しては、本稿は日本語を対象としたスコープ解析の第一歩であるため、まずは頻度が高くかつ基本的な現象である、文末に最も近い事象（主事象）とそれに隣接する事象（従属事象）について、従属事象が主事象中の否定、推量、疑問という機能表現のスコープ内にあるかを判定する課題に取り組んだ。2のデータに関しては、BCCWJのウェブドメインであるYahoo!知恵袋を対象に、人手でスコープ内外の付与を行った。本研究では、曖昧性が高く難しい部分であることから、述語間が「テ形接続」である場合に限定してサンプリングを行った結果、否定2,178文、推量501文、疑問910文の計3,589文に対してスコープを付与することができた。

スコープを付与したデータを用いて、機械学習による二値分類を行うスコープ解析器を構築した。交差検定およびオープンテストによってスコープ解析器の性能を測ったところ、交差検定で86.8%、オープンテストで79.7%の正解率が得られた。素性については検討の余地が残されており、今後の重要な研究課題である。

6 結論

事実性解析には、事象に含まれる機能表現、疑問詞を含む副詞、文節境界を越えて事実性に影響を与える語とそのスコープ、その他の4種類の問題が含まれている。それぞれは単独でも一つの研究課題になるほどに、容易な問題ではないが、事実性解析ではさらにその組み合わせがあるため、性能の向上が難しい。本研究では、事実性解析の課題分析を行うために、機能表現のみを用いたルールベースの事実性解析器を構築し、1,533文に含まれる3,734事象に適用した結果の誤りを分析した。このとき全ての事象表現について、述語に続く機能表現に対して意味ラベルを付与した。

主事象の事実性解析については、機能表現の意味ラベルが正しく解析できれば、現在の意味ラベルの体系と本研究で用いた単純な規則だけでも、90%に近い正解率が得られることがわかった。本研究で用いた規則は人手で構築したものであるため、その整備は必要ではあるものの、それよりもむしろ、現在の機能表現の意味ラベル体系に基づいて機能表現解析モデルの研究開発を行うことに一定の支持を与えるものと考えられる。また、機能表現解析の問題を除けば、誤りの半数は副詞に起因するものであった。したがって、事実性解析は副詞の意味解析の研究を動機付ける良い課題となりうる。

従属事象の事実性解析は、主事象に比べて考慮すべき要素が多く、性能も低い。従属事象でのみ考慮すべき要素は大きく二つあり、文節境界を越えて事実性に影響を与える述語と、従属事象に直接付随しない機能表現の影響である。文節境界を越えて事実性に影響を与える述語については、既存の事象選択述語辞書が一定のカバレッジを持っており、これを利用することで誤りの多くを解消できる可能性がある。しかし、複合語のカバレッジに問題があるなど、こうしたリソースの整備が今後の課題であることがわかった。

従属事象に直接付随しない機能表現については、直接の親の事象に付随する機能表現の影響を受ける可能性があるが、その他の事象表現に付随する機能表現の影響はほとんど無視できることも明らかになった。前者の場合については、誤りの半分近く(36/80)にあたる「直接の項」は我々の分析データを見る限り、全ての場合においてスコープ内に来るので、述語項構造解析の結果に基づいてスコ

プを広げることにより，事実性解析の性能を向上させることができる．一方で，誤りのうち4割以上(33/80)「テ形接続」「項を修飾」「名詞述語を修飾」等の接続パターンの場合には，スコープ内外の選択が高度に曖昧であり，これらのパターンのスコープを決定する課題に注力することに一定の効用があることがわかる．それ以外の主要な接続パターンはスコープの範囲を規則ベースで決めても大きな問題は生じそうにない．また，離れた事象対と比較して，隣接事象対のスコープを特定する方が，事実性解析に対して大きなゲインが期待できる．実際にスコープを手で付与し，事実性解析に取り入れることで，CT+以外の性能，特に再現率を向上させることができた．このことから，隣接事象対のスコープ判定を精緻に行うことが，事実性解析の性能向上に貢献することを確認できた．

次に，従属事象において課題とされたスコープ解析課題のうち，スコープ内外が曖昧である「テ形接続」について，スコープの自動解析を行った．スコープ解析については，日本語を対象としたスコープ解析の課題設計が自明ではなく，スコープ解析に利用可能なデータも存在しない，という課題があった．そのため本研究では，まず事実性解析への応用に適合する問題設計を行い，スコープ情報を付与したコーパスを新たに構築した．具体的には，主事象と間近の従属事象が「テ形接続」になっている事例について，スコープの内外を付与したコーパスを構築したところ，否定2,178文，推量501文，疑問910文の計3,589文に対してスコープを付与することができた．構築したコーパスを用いて機械学習に基づくスコープ解析器を構築したところ，79.7%の正解率を実現したが，素性については検討の余地が残されており，今後の重要な研究課題である．

本研究で報告した誤り分析・課題分析は「Yahoo!知恵袋」のコーパスを用いており，他のドメインやスタイルの文章で同様の傾向が得られるかは明らかでない．今後は調査の範囲を広げ，問題の性質の一般化を図る．また，自動解析性能の向上については，更新ルールや辞書の整備も必要な課題ではあるものの，本研究では正解を与えた，機能表現の意味ラベルを自動で解析する課題に注力することが重要であると考えられる．スコープ解析においては「て接続」に注力したが，その他の接続表現においても，スコープ内外を判定することが有効である事例（スコープ内外の偏りが比較的大きくない事例）についても取り組む必要がある．

謝辞

本研究を進めるにあたって、多くの方にご協力をいただきました。ここに、心より感謝の意を表します。

主指導教官である乾健太郎教授には、お忙しい中、研究活動全般にわたり、終始温かいご指導、ご助言をいただきました。心より感謝を申し上げます。ご多忙の中、審査委員をお引受け下さり、便宜をはかってくださった、篠原歩教授、木下哲男教授に深く感謝致します。本研究を進めるにあたって、NICTの水野淳太研究員には、数々の的確なご助言をいただき、細部まで面倒をみていただきました。心より感謝致します。また、種々のアノテーションを行うにあたって仕様を議論するとともに、実際の作業を非常に丁寧に行ってくださった、福原裕一研究員、菅野美和さんに深く感謝致します。山梨大学の松吉俊助教には、拡張モダリティのアノテーションに関して、丁寧に教えていただきました。心より感謝致します。本研究を進めるにあたり、岡崎直観准教授、松林優一郎特任助教、井之上直也助教、上岡裕大君をはじめ、乾・岡崎研究室の皆様からは、様々なご助言をいただき、相談にのっていただくとともに、研究生生活を暖かく支えていただきました。心より感謝を申し上げます。

少し遠回りをしてしまいましたが、無事に博士論文を提出するに至ることができました。これも、乾・岡崎研究室の皆様や、学会で知り合った皆様、大学生活において関わってくださった皆様の支えがあってこそです。何一つ欠けてはここまで辿りつけなかったと思います。最後になりましたが、研究生生活や大学生活において関わってくださった全ての皆様に深く感謝致します。本当にありがとうございました。

参考文献

- [1] Lauri Karttunen and Annie Zaenen. Veridicity. In Graham Katz, James Pustejovsky, and Frank Schilder, editors, *Annotating, Extracting and Reasoning about Time and Events*, No. 05151 in Dagstuhl Seminar Proceedings, 2005.
- [2] Roser Saurí and James Pustejovsky. Determining modality and factuality for text entailment. In *Proceedings of the International Conference on Semantic Computing*, pp. 509–516, 2007.
- [3] Andrew Hickl. Using discourse commitments to recognize textual entailment. In *Proceedings of the 22nd International Conference on Computational Linguistics*, Vol. 1, pp. 337–344, 2008.
- [4] Roser Saurí and James Pustejovsky. FactBank: a corpus annotated with event factuality. *Language resources and evaluation*, Vol. 43, No. 3, pp. 227–268, 2009.
- [5] 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.
- [6] 川添愛, 齊藤学, 片岡喜代子, 崔栄殊, 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.
- [7] 川添愛, 齊藤学, 片岡喜代子, 崔栄殊, 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.

- [8] 梅澤俊之, 西尾華織, 松田源立, 原田実. 意味解析システム sage の精度向上とモダリティの付与と辞書更新支援系の開発. 言語処理学会第 14 回年次大会発表論文集, pp. 548–551, 2008.
- [9] 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.
- [10] Charles Fillmore. The Case for Case. In Emmon Bach and Robert T Harms, editors, *Universals in Linguistic Theory*, Vol. 25, pp. 1–88. Holt, Rinehart and Winston, 1968.
- [11] 益岡隆志. 日本語モダリティ探求. くろしお出版, 2007.
- [12] 日本語記述文法研究会 (編). 現代日本語文法 4. くろしお出版, 2003.
- [13] 日本語記述文法研究会 (編). 現代日本語文法 3. くろしお出版, 2003.
- [14] Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. The penn discourse treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, pp. 2961–2968, 2008.
- [15] Roser Saurí. FactBank 1.0 Annotation Guidelines, 2008.
- [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] Roser Saurí, Jessica Littman, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky. TimeML Annotation Guidelines, Version 1.2.1, 2006.
- [18] Marie-Catherine de Marneffe, Christopher D. Manning, and Christopher Potts. Did it happen? The pragmatic complexity of veridicality assessment. *Computational Linguistics*, Vol. 38, No. 2, pp. 301–333, 2012.

- [19] L. R Horn. Metalinguistic negation and pragmatic ambiguity. *Language*, Vol. 61, No. 1, pp. 128–174, 1985.
- [20] 田中リベカ, 戸次大介, 川添愛. MCN コーパス : ガイドライン設計とその運用. 言語処理学会第 19 回年次大会発表論文集, pp. 77–80, 2013.
- [21] 田中リベカ, 川添愛, 戸次大介. MCN コーパス : 「ノダ」にみるガイドライン作成の手法. 言語処理学会第 20 回年次大会発表論文集, pp. 1071–1074, 2014.
- [22] 宇津木舞香, 佐藤未歩, 青木花純, 田中リベカ, 川添愛, 戸次大介. MCN コーパスにおける形式名詞「はず」「わけ」「つもり」のアノテーション. 言語処理学会第 20 回年次大会発表論文集, pp. 1067–1070, 2014.
- [23] 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.
- [24] 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 *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 01*, pp. 314–321, 2008.
- [25] Roser Saurí and James Pustejovsky. Are you sure that this happened? assessing the factuality degree of events in text. *Computational Linguistics*, Vol. 38, No. 2, pp. 261–299, 2012.
- [26] Marc Light, Xin Ying Qiu, and Padmini Srinivasan. The language of bio-science: Facts, speculations, and statements in between. In *Proceedings of BioLink 2004 workshop on linking biological literature, ontologies and databases*, pp. 17–24, 2004.

- [27] 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.
- [28] 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.
- [29] 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, No. 5, pp. 123–146, 2007.
- [30] 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.
- [31] 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.
- [32] Yudai Kamioka, Kazuya Narita, Junta Mizuno, Miwa Kanno, and Kentaro Inui. Semantic annotation of Japanese functional expressions and its impact on factuality analysis. In *Proceedings of The 9th Linguistic Annotation Workshop*, pp. 52–61, June 2015.
- [33] 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. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning — Shared Task*, pp. 1–12, 2010.
- [34] Roser Morante and Eduardo Blanco. *SEM 2012 shared task: Resolving the scope and focus of negation. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics - Volume 1: Proceedings of the main conference and the shared task*, pp. 265–274, 2012.

- [35] 松吉俊. 否定の焦点情報アノテーション. 自然言語処理, Vol. 21, No. 2, pp. 249–270, 2014.
- [36] 飛田良文, 浅田秀子. 現代副詞用法辞典. 東京堂出版, 1994.
- [37] 森田良行, 松木正恵. 日本語表現文型 用例中心・複合辞の意味と用法. アルク, 1989.
- [38] 遠藤織枝, 小林賢次, 三井昭子, 村木新次郎, 吉沢靖 (編). 使い方の分かる類語例解辞典 新装版. 小学館, 2003.
- [39] 水野淳太, 渡邊陽太郎, エリックニコルズ, 村上浩司, 乾健太郎, 松本裕治. 文間関係認識に基づく賛成・反対意見の俯瞰. 情報処理学会論文誌, Vol. 52, No. 12, pp. 3408–3422, 2011.
- [40] 後藤淳, 大竹清敬, Stijn De Saeger, 橋本力, Julien Kloetzer, 川田拓也, 鳥澤健太郎. 質問応答に基づく対災害情報分析システム. 自然言語処理, Vol. 20, No. 3, pp. 367–404, 2013.
- [41] 工藤拓, 松本裕治. チャンキングの段階適用による日本語係り受け解析. 情報処理学会論文誌, Vol. 43, No. 6, pp. 1834–1842, 2002.
- [42] Erik F. Tjong Kim Sang. Noun phrase recognition by system combination. In *Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference*, NAACL 2000, pp. 50–55, 2000.
- [43] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, Vol. 20, No. 1, pp. 37–46, 1960.
- [44] John D. Lafferty, Andrew McCallum, and Fernando C. N. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, pp. 282–289, 2001.

- [45] Naoaki Okazaki. CRFsuite: a fast implementation of Conditional Random Fields (CRFs), 2007.
- [46] 水野淳太, 成田和弥, 乾健太郎, 大竹清敬, 鳥澤健太郎. 拡張モダリティ解析器の試作と課題分析. ALAGIN & NLP 若手の会 合同シンポジウム, 2013.
- [47] Charles Sutton, Andrew McCallum, and Khashayar Rohanimanesh. Dynamic conditional random fields: Factorized probabilistic models for labeling and segmenting sequence data. *The Journal of Machine Learning Research*, Vol. 8, pp. 693–723, May 2007.
- [48] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, Vol. 9, pp. 1871–1874, 2008.
- [49] 国立国語研究所. 分類語彙表. 大日本図書, 2004.
- [50] 南不二男. 現代日本語の構造. 大修館書店, 1974.
- [51] 有田節子. 日本語条件文と時制節性. くろしお出版, 2007.
- [52] 田窪行則. 日本語の構造 推論と知識管理. くろしお出版, 2010.
- [53] Junta Mizuno, Canasai Kruengkrai, Kiyonori Ohtake, Chikara Hashimoto, Kentaro Torisawa, Julien Kloetzer, and Kentaro Inui. Recognizing complex negation on Twitter. In *The 29th Pacific Asia Conference on Language, Information and Computation (PACLIC-29)*, pp. 544–552, 2015.
- [54] Naoaki Okazaki. Classias: a collection of machine-learning algorithms for classification, 2009.

発表文献一覧

受賞一覧

1. 情報処理学会第77回全国大会 学生奨励賞 (第一著者である上岡裕大が受賞)
2. 情報処理学会第77回全国大会 大会奨励賞 (第一著者である上岡裕大が受賞)

学術論文誌

1. 成田和弥, 水野淳太, 上岡裕大, 菅野美和, 乾健太郎. 誤り分析に基づく日本語事実性解析の課題抽出. 自然言語処理, Vol.22 No.5, pp. 397-432, December 2015.

国際会議論文

1. Yudai Kamioka, Kazuya Narita, Junta Mizuno, Miwa Kanno and Kentaro Inui. Semantic Annotation of Japanese Functional Expressions and its Impact on Factuality Analysis. In Proceedings of the 9th Linguistic Annotation Workshop (LAW IX 2015), June 2015.
2. Kazuya Narita, Junta Mizuno and Kentaro Inui. A Lexicon-based Investigation of Research Issues in Japanese Factuality Analysis. In Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP 2013), pp.587-595, October 2013.
3. Kazuya Narita, Tetsuya Sakai, Zhicheng Dou and Young-In Song. MSRA at NTCIR-10 1CLICK-2. In Proceedings of the 10th NTCIR Conference, pp.218-224, June 2013.

国内会議・研究会論文

1. 成田和弥, 水野淳太, 上岡裕大, 菅野美和, 乾健太郎. 機能表現に基づく日本語事実性解析. 言語処理学会第 21 回年次大会予稿集, pp.1032-1035. March 2015.
2. 上岡裕大, 成田和弥, 菅野美和, 水野淳太, 乾健太郎. 日本語文における機能表現意味ラベル付与と事実性解析への効果. 情報処理学会第 77 回年次大会予稿集, pp. 221-222, March 2015. (学生奨励賞, 大会奨励賞)
3. 上岡裕大, 菅野美和, 成田和弥, 水野淳太, 乾健太郎. 述部機能表現に対する意味ラベル付与と事実性解析への適用. NLP 若手の会 第 9 回シンポジウム, September 2014.
4. 上岡裕大, 成田和弥, 水野淳太, 乾健太郎. 述部機能表現に対する意味ラベル付与. 情報処理学会 第 216 回自然言語処理研究会・第 101 回 音声言語情報処理研究会, Vol.2014-NL-216/Vol.2014-SLP-101, No.9, pp.1-9, May 2014.
5. 成田和弥, 水野淳太, 乾健太郎. 事実性解析における否定のスコープ解析. ALAGIN & NLP 若手の会 合同シンポジウム, September 2013.
6. 水野淳太, 成田和弥, 乾健太郎, 大竹清敬, 鳥澤健太郎. 拡張モダリティ解析器の試作と課題分析. ALAGIN & NLP 若手の会 合同シンポジウム, September 2013.
7. 成田和弥, 水野淳太, 乾健太郎. 語彙知識と構成性に基づく日本語事実性解析. 言語処理学会第 19 回年次大会予稿集, pp.98-101, March 2013.
8. 成田和弥, 水野淳太, 乾健太郎. 日本語事実性解析課題の経験的分析. 情報処理学会研究報告 第 204 回自然言語処理研究会, Vol.2011-NL-204 No.17, pp.1-8, November 2011.
9. 成田和弥, 水野淳太, 乾健太郎. 日本語事実性解析課題の経験的分析. NLP 若手の会 第 6 回シンポジウム, September 2011.

公開資源

1. 機能表現タグ付与コーパス <http://www.cl.ecei.tohoku.ac.jp/index.php?Open%20Resources%2FJapanese%20FE%20Corpus>
2. 事象選択述語辞書 <http://www.cl.ecei.tohoku.ac.jp/index.php?Open%20Resources%2FJapanese%20ESP%20Dictionary>
3. 日本語拡張モダリティ解析器 Zunda <https://github.com/jmizuno/zunda>