

B4IM2018

修士論文

日本語述部機能表現の解析と モダリティ解析への適用

上岡裕大

2016年2月16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
修士 (情報科学) 授与の要件として提出した修士論文である。

上岡裕大

審査委員：

乾 健太郎 教授 (主指導教員)

田中 和之 教授

伊藤 彰則 教授

岡崎 直観 准教授 (副指導教員)

日本語述部機能表現の解析と モダリティ解析への適用*

上岡裕大

内容梗概

自然言語処理において内容語に関する解析は広く行われている一方で、機能語の扱いは十分であるとは言い難い。これまでにいくつかの研究によって機能表現を扱う言語資源が整備されたが、実際の日本語文に対するアノテーションの考察や、応用タスクに適用した際の分析などを考慮すると、未だに改善の余地がある。拡張モダリティ解析を主眼に置いた機能表現解析を目的とし、日本語機能表現の意味体系の設計と現代日本語書き言葉均衡コーパスに対するラベル付与を行った。本論文では、言語処理の観点からコーパス構築における問題点と、構築したコーパスを用いた機能表現解析の現状を報告する。また、機能表現解析結果を拡張モダリティ解析に適用し、機能表現解析の寄与と問題点についても議論する。

キーワード

機能表現，コーパスアノテーション，モダリティ，意味解析

*東北大学 大学院情報科学研究科 システム情報科学専攻 修士論文, B4IM2018, 2016年2月16日.

Semantic Labeling of Japanese Functional Expressions and its Impacts on Extended Modality Analysis*

Yudai Kamioka

Abstract

Recognizing the meaning of functional expressions is essential for natural language understanding. This is a difficult task, owing to the lack of a sufficient corpus for machine learning and evaluation. In this study, we design a new annotation scheme and construct a corpus containing 5,736 Japanese sentences and 20,488 functional expressions. Our scheme achieves high inter-annotator agreement with kappa score of 0.85. In the experiments, we confirmed that machine learning-based functional expression analysis achieves F-score of 88.10, and contributes to Extended modality analysis.

Keywords:

Functional Expressions, Corpus Annotation, Modality, Semantic Parsing

*Master's Thesis, System Information Sciences, Graduate School of Information Sciences, Tohoku University, B4IM2018, February 16, 2016.

目次

1	はじめに	1
1.1	研究課題	1
1.2	本研究の取り組み	3
1.3	本論文の構成	4
2	関連研究	5
2.1	機能表現に関する言語資源の研究	5
2.2	モダリティに関する研究	6
3	機能表現意味ラベル付与コーパス	8
3.1	機能表現	8
3.2	機能表現意味体系の設計	10
3.3	機能表現意味ラベル付与コーパスの構築	12
3.3.1	意味ラベルの表現方法	12
3.3.2	アノテーション作業	13
3.4	機能表現意味ラベル付与コーパス	14
3.5	作業間一致率	17
4	機能表現解析	22
4.1	条件付き確率場	22
4.2	評価実験	23
4.2.1	学習素性	23
4.2.2	Baseline: 最長一致法	24
4.3	評価	25
4.3.1	評価尺度	25
4.3.2	評価結果	25
4.4	議論	26
5	拡張モダリティ解析	30

5.1	拡張モダリティタグ	30
5.1.1	ロジスティック回帰による拡張モダリティ解析	32
5.2	実験	32
5.3	評価結果	34
5.3.1	学習素性の比較	34
5.3.2	意味体系の比較	34
5.3.3	意味ラベルの付与方法の比較	34
5.3.4	項目ごとの評価結果	35
6	おわりに	36
	謝辞	37

目 次

1	Graphical model of linear-chain CRF.	23
---	--	----

表 目 次

1	意味ラベルの付与例	13
2	作業者が判断に迷った事例	14
3	機能表現意味ラベルコーパスの統計情報	15
4	構成する形態素数別の複合辞数	16
5	事象に連続して付属する機能表現数	17
6	作業者間一致率の κ 値	18
7	意味ラベルの作業者間一致率	21
8	学習素性の一覧	24
9	機能表現解析器の評価結果 (Closed)	26
10	機能表現解析器の評価結果 (Open)	26
11	意味ラベル別の評価結果	28
12	誤り分析	29
13	機能表現の長さ別の評価結果	30
14	拡張モダリティタグの一覧	31
15	拡張モダリティ解析の学習素性	33
16	4 項組の評価結果	33
17	最頻出のタグを除く項目ごとの評価結果	35

1 はじめに

日本語文では、述語や命題に対して機能的な語が付属することによって、様々な情報が付加される。例えば、(1a)は、述語「降る」に対して助動詞「ない」が付属し、述語が表す「降る」という事象が起こらないことを表している。また、(1b)は、「明日は雨が降る」に対して「かもしれない」が付属することで、話者による不確かな判断であるという情報が付加される。

(1) a. 明日は雨が [降ら] ない。

b. 明日は雨が [降る] かもしれない。

このように述部に対して何らかの情報を付加する機能的な語は機能表現と呼ばれる。また、本研究では、機能表現が付属することによって付加される情報を機能表現の意味と呼ぶ。機能表現は、大きく二つに分類できる。一つは、(1a)のような助詞・助動詞などの機能語である。もう一つは、複数の語が複合して、ひとまとまりの形で辞的な機能を果たす表現であり、これを複合辞と呼ぶ。例えば、(1b)における「かもしれない」は、IPA品詞体系に従えば、「かも」「しれ」「ない」の3つの形態素から構成されており、「かもしれない」をひとつの表現、すなわち、複合辞として不確かな判断の意味を表す。

これらの機能表現を抽出し、その意味を認識することは、幅広いタスクにおいて重要な課題である。例えば、上記の例のように述語に付属する機能表現を正しく認識することによって、述語が表す事象が成立したか否か、あるいは、成立が望まれているか否かといった情報を抽出することができる。

1.1 研究課題

しかし、機能表現の種類は多岐にわたり、否定を表す機能表現を例に挙げると、「ない」「ではない」「わけない」「はずがない」などの表現がある。これらの機能表現を応用タスクごとに認識するコストは大きい。そこで本研究では、機能表現を意味単位で抽象化することで、応用タスクの解析性能が向上するという仮説を

立てた．例えば，先の否定表現のそれぞれにあらかじめ「否定」という意味ラベルを付与しておけば，応用タスクではこの意味ラベルを用いて抽象化された単位で処理することができる．本論文では，文中から機能表現を抽出し，意味ラベルを付与するタスクを機能表現解析と定義し，機能表現解析の現状を報告する．

機能表現解析を行うにあたって，以下の2つの課題がある：

(A) 学習・評価に必要なコーパスが一般に利用可能ではない

(B) 既存の意味体系は，表現および意味のカバレッジが不足している

(A) について，これまでに機能表現を扱った研究には [1, 2] などがある．これらの研究では，人手で作成した正解事例をもとに統計的学習手法を用いた解析モデルにより機能表現の抽出を行った．しかし，これらの研究において作成されたコーパスは公開されておらず，我々の知る限りでは，機能表現解析に必要なコーパスとして一般に利用可能なものが存在しない．

(B) について，利用可能な意味体系として，日本語機能表現辞書「つつじ」[3] において定義される意味体系がある．つつじは日本語機能表現を階層的に収録した電子的に利用可能な辞書であり，先の [1, 2] や [4] においても，本体系を利用している．しかし，つつじを利用した研究において，意味ラベルやエントリを追加して利用する必要があるなど，機能表現意味体系の不適合やカバレッジ不足が指摘されている．こうした問題は，機能表現の捉え方が応用タスクに依存するという性質に起因するものである．例えば，文 (2a)(2b) は，いずれも事象「降る」に対する話者の推測を表しているが，2つの文の間には話者による確信の度合いに違いがあると考えられる．(2a) の文よりも (2b) の文のほうが話者が事象に対して強い確信を持っていると言えるだろう．話者による確信の度合いを精緻に認識したいタスクにおいては，これらは別の機能表現として異なる意味ラベルを与える必要がある．

(2) a. 明日は雨が [降る] かもしれない .

b. 明日は雨が [降る] に違いない .

1.2 本研究の取り組み

これらの問題を踏まえ、本研究では、新たに機能表現の意味体系を設計し、一般に利用可能な機能表現意味ラベル付与コーパスを構築した。まず、(B)の問題を解決するために、応用タスクを設定した上で意味体系の設計を試みた。応用タスクとして、拡張モダリティ解析を採用する。拡張モダリティ解析の部分課題である事実性解析について、ルールベースの解析器を構築し、その解析結果をフィードバックしながら意味体系を設計し、72種類の意味ラベルの定義し、それぞれの付与対象となる表現を記載したアノテーションガイドラインを作成した。次に、(A)の問題に対する解決策として、『現代書き言葉均衡コーパス (BCCWJ)』に対して、設計したガイドラインに基づいて人手で意味ラベルを付与した機能表現意味ラベル付与コーパスを構築した。本コーパスは、5,346文の訓練データと300文のテストデータからなり、作業者間一致率として、 ≈ 0.85 を達成した。また、コーパスはガイドラインと合わせて、一般に利用可能なデータとしてウェブ上で公開している。

上記の問題を解決した上で、機能表現解析を系列ラベリング問題として定式化し、構築したコーパスを用いて、条件付き確率場 (CRF) を用いて機能表現解析を行った。機能表現解析性能として88.10%のF値を達成し、辞書ベースの最長一致法と比較して機械学習による解析が有効であることを示すとともに、コーパスが統計的学習を行うのに十分な一貫性があることを示した。さらに、[3]で定義される拡張モダリティの解析において、本研究の機能表現解析結果を学習素性として用いると、機能表現解析結果を用いない場合に比べて性能が向上することを検証した。

本研究の貢献は、以下の通りである。

1. 拡張モダリティ解析への応用を考慮した機能表現の意味体系を設計し、アノテーションガイドラインと機能表現意味ラベル付与コーパス、および解析器を公開した。
2. CRFによる機能表現解析を行い、88.10%のF値を達成した。
3. 機能表現解析結果を拡張モダリティ解析に適用し、機能表現解析が拡張モ

ダリティ解析に貢献することを示した。

1.3 本論文の構成

次節以降では、まず、2節で機能表現および機能表現を用いた拡張モダリティ解析に関する先行研究を概観し、3節で作成した機能表現の意味体系およびコーパス構築について議論する。4節でコーパスを用いた機能表現解析結果を報告し、5節でモダリティ解析における機能表現解析の貢献を示す。最後に6節でまとめと今後の課題を述べる。

2 関連研究

2.1 機能表現に関する言語資源の研究

日本語の機能表現に関しては、データベースや辞書の構築、範囲の検出、意味の自動認識など、さまざまな研究がされてきた。機能表現に関する日本語の言語資源として、日本語複合辞用例データベースがある [5]。用例データベースは、現代語複合辞用例集 [6] を元に、新聞記事に含まれる 337 種類の複合辞候補に対して内容的用法と機能的用法の別をアノテートしたデータベースであり、各項目ごとに最大 50 件の用例データを収録している。また、大規模な数の機能表現を処理する基礎として、松吉らによる日本語機能表現辞書『つつじ』 [3] がある。つつじでは、言語学的文献を参考にして得た見出し語 341 件について、種々の異形を考慮し、約 17,000 種類の機能表現を収録している。つつじの見出し体系は 9 つの階層を持つ階層構造となっており、見出し語・意味・派生・機能語の交替・音韻的变化・とりたて詞の挿入・活用・「です/ます」の有無・表記のゆれなどを考慮している。収録される機能表現は、既存の機能表現リストとの比較により、各々の見出し語に対してほぼすべての異形を網羅しているとされる。用例データベースは新聞記事を対象にしているため、ウェブテキストのような崩れた表現を収録しておらず、また、機能表現の意味はアノテートされていない。つつじは、機能表現を意味カテゴリに分類し、異形も網羅的に収録しているが、機能表現のカバレッジ不足を指摘する研究もある [1, 4]。これは、つつじが文献を参考に設計されたことや、応用タスクを明確に決定していないことが原因となり、機能表現の意味の粒度が適切でないためであると考えられる。また、実際につつじで定義される機能表現とその意味を手で付与したデータは我々の知る限り一般に利用可能でない。英語においては、いくつかのコーパスが公開されている。Szarvas らは、医療テキスト中の否定表現、推量表現およびそのスコープをアノテーションした BioScope を構築した [7]。

機能表現の範囲の検出に関する研究としては、[8, 9, 10] がある。[8] は、機能表現検出を形態素を単位とするチャンキング問題として定式化し、形態素解析結果から機械学習によって機能表現を検出した。[9, 10] は、つつじを用いた機能表

現解析を行った。特に [10] は、大規模な均衡コーパスである『現代日本語書き言葉均衡コーパス』において、機能的用法・内容的用法の曖昧性を持つ機能表現を対象として、機械学習により用法判定を行う手法を適用し、その性能を評価した結果を報告した。当文献における機能表現とは、我々の扱う機能表現のうち複合辞に相当するものである。複合辞となり得る表記を構成する形態素が、機能語となるのか複合辞の部分となるのかという曖昧性解消に取り組んだ。条件付確率場を利用したチャンキングを用いた実験の結果、97%近い F 値を達成している。しかしながら、彼らの研究はいずれも、機能的用法、内容的用法の区別に留まっており、機能表現の意味を特定するまでには至っていない。機能表現に関する意味の自動認識に関する研究には、今村らの意味ラベルタガー [1] がある。今村らは、形態素解析結果に対して、述部を同定し、つつじの意味ラベルを機能表現に付与する意味ラベルタガーを構築した。機能表現辞書つつじと識別モデルに基づく最尤選択を組み合わせて機能表現を同定しており、辞書を用いてラティスを作成することにより、異なるドメイン間において系列ラベリングより高い精度で意味ラベルを付与することに成功した。しかし、タグ付けに用いた学習・評価用のコーパスの公開や、タグ付けにおける課題などは明らかにされていない。機能表現の検出だけでなく、意味も含めて解析する研究は、我々が調査した限り、今村ら以外にない。

2.2 モダリティに関する研究

機能表現解析の応用として、係り受け解析 [11] や機械翻訳 [12, 13] などへ応用された例がある。英語圏では、[14] が、modal words のリストと no や any などの否定的な文脈を表す言語マーカールのリストを用いて自動的にイベントの事実性を分類した。その他の応用として、[15] は、n't などの否定表現を矛盾認識の学習に用いている他、[16] は、構造ベースのモダリティ解析によって機械翻訳の性能が向上することを示している。

本研究では、機能表現解析をモダリティ解析に適用する。[4] による事実性解析は、モダリティ解析の部分技術である。事実性とは、文中の事象が現実が発生したできごとかどうかに関する筆者の判断に関する情報である。例えば、(3a) では、

「それが故障である」ことが、実際には起こっていないことであることを述べており、(3b)では、「彼が先に帰っている」ことについて、その可能性が高いことを述べている。このとき、(3a)の「ん」が否定を表す機能表現であること、(3b)の「かもしれない」が推量を表す機能表現であることが自動的に認識できると、事実性解析のための大きな手がかりとなる。

(3) a. それは故障ではありません。

b. 彼は先に帰っている かもしれない。

[4]は、日本語に対するルールベースの事実性解析器を構築し、それを拡張モダリティ付与コーパス [17]の一部に適用した。[4]の解析器は、[18]の英語における解析器を、日本語に適応させたものであり、機能表現のような、事実性解析の手がかりとなる表現に基づく、事実性の更新ルールを構成的に組み合わせることで、決定的に事実性を解析するモデルである。その誤り分析を行うことで、機能表現の曖昧性解消や、否定・推量表現のスコープ解析が事実性解析において重要であることを示している。

3 機能表現意味ラベル付与コーパス

3.1 機能表現

コーパスを構築するにあたり，ここで本研究における機能表現を定義する．日本語文には，助詞や助動詞などの機能語と，複数の語が組み合わせることによって助詞や助動詞相当の役割を果たしたり，モダリティ相当の意味を持つ表現があり，これらをまとめて機能表現と呼ぶ．本論文では，応用が拡張モダリティ解析であることから，述部の機能表現を主な解析対象とする．以下に機能表現の例を示す．

- (4) a. きれいな [雪] だ わ .
b. 明日は雨が [降る] そ う だ .
c. 十六世紀のキリスト教は、主としてポルトガルの [宣教師] に よ っ て も た ら さ れ た .
d. 二つに一つは [助から] な い か も し れ な い と 思 っ て い た の だ が …

(4a) の下線部はそれぞれ断定の意味を表す助動詞，および感動の意味を表す終助詞であり，いずれも個別の機能表現である．(4b) の下線部は伝聞の意味を表す助動詞であり，本研究で扱う機能表現である．一方で，(4b) 中の「は」や「が」は格助詞，すなわち機能表現であると言えるが，述語に付随しないため，本研究では解析の対象としない．

また，(4c) の下線部は，格助詞「に」，動詞「よる」，接続助詞「て」との連語，(4d) 下線部は，副助詞「か」，係助詞「も」，動詞「しる」，助動詞「ない」との連語である．(4c) および (4d) の例は，「格助詞と動詞と接続助詞の連語」というように文法的に説明することができるが，意味を解釈する上では，文法的な構成要素の組み合わせとして考えるよりも，全体としてひとつの役割を担う表現として解析するほうが後の処理において扱いやすい．これらの表現は，ひとまとまりの辞として複合辞¹と呼ばれる．すなわち，(4c) は，後続の動詞「もたらず」

¹同様の考え方を初めに導入した永野氏に倣い「複合辞」と呼称するが「複合助辞」「助詞相当連語・助動詞相当連語」という名称が用いられることもある [19][20] ．

の主体を示す格助詞相当の複合辞であり、(4d) 下線部は、未来のことについての不確実な判断を表すモダリティ相当の複合辞であると考える。

どのような表現を複合辞とするかの判定については、松木の研究 [20] において考察されている。松木は、認定基準として、

第1種複合辞

- 1 形式的にも意味的にも辞的な機能を果たしていること
- 2 形式全体として、個々の構成要素の合計以上の独自の意味が生じていること

第2種・第3種複合辞

- 1 形式的にも意味的にも辞的な機能を果たしていること
- 2 中心となる「詞」は実質的意味が薄れ、形式的・関係構成的に機能していること
- 3 2の語に他の辞的な要素等が結合して一形式を構成する場合、その要素の持つ意味が2の語に単に付加されたものではなく、形式全体として独自の意味が生じていること

と設定している。第1種複合辞は助詞・助動詞のみが複合して出来た複合辞、第2種・第3種はそれぞれ形式名詞・形式用言を含む複合辞を指す。[20] 内に挙げられる複合辞の例を以下に示す。(5a) は機能語の助動詞と助詞からなる複合辞であり、(5b)、(5c) は、それぞれ名詞「ところ」や動詞「来る」を含む複合辞である。

- (5) a. 僕は小学校を卒業したばかりで十五歳、月を数えると十三歳何ヶ月という頃、…。
- b. この問題について、いろいろ説明したところで、理解してもらえないに違いない。
- c. わたしたちときたら、部屋を見せてもらったとき、鍵の有無なんて、まったく意識にのぼりませんでした。

本研究における複合辞の認定基準は，松木の認定基準を参考にした．この基準を採用する理由は，複合辞の認定基準に加えて，どれだけ複合辞らしいかという複合辞性の概念を取り入れている点にある．松木らの基準では，複合辞性の尺度として，(i) 構成要素の緊密化の度合い，(ii) 形式名詞・形式用言の形式化の度合い，(iii) 形式用言の文法範疇喪失の度合いの3項目を挙げている．本研究でも複合辞の認定時にこれらの尺度を考慮に入れた．ただし，いずれの尺度についても境界は曖昧であり，厳密な境界を設けることはせず，事例ごとに判定を行った．

3.2 機能表現意味体系の設計

機能表現に意味ラベルを付与するために機能表現を72種類の意味カテゴリに分類した．本節では，機能表現の意味体系の設計について述べる．

意味の粒度をどの程度にすべきかの判断は難しい．本研究では，応用タスクで特別な処理をせず利用できるような粒度の小さい意味体系の設計を試みた．また，同様の理由で，複合辞は積極的に認めるという方針をとった．本論文では，解析の際に機能表現の支配度が高いという仮説に基づき，事象の拡張モダリティ[17]を自動推定する拡張モダリティ解析を応用タスクに設定し，モダリティ表現を中心に意味体系の設計を進めた．特に，拡張モダリティの真偽判断に相当する事実性に関して，日本語文では文末の事象の事実性は機能表現のみによって決定できる場合が多く，機能表現が及ぼす影響を検証しやすい．そこで，ルールに基づき真偽判断を決定的に分類する[4]の事実性解析システムの出力結果をフィードバックしながら意味体系を設計した．具体的には，つつじ[3]やその参考文献[21]を元にはじめの仕様を定め，アノテーション仕様にしたがってコーパスを構築し，機能表現の意味ラベルが付与されたデータを用いて，機能表現に基づく事実性解析を行い，誤分類された事例を中心に機能表現の付与基準を検討し，アノテーション仕様を改善するという手順を数回繰り返した．

最終的に，72種類の意味ラベルを定義した．つつじの意味体系では，機能表現を89種類の意味カテゴリに分類している．本研究では，述部の機能表現のみを対象にしているため，意味ラベルの種類はつつじより少ない．例えば，「不确实」「高确实性」という意味ラベルは本研究で新たに定義する意味ラベルである．

(6) a. 明日は雨が [降る] かもしれない不確実 .

b. 明日は雨が [降る] に違いない高確実性 .

文(6a)(6b)は、いずれも事象「降る」に対する話者の推測を表しているが、2つの文の間には話者による確信の度合いに違いがあると考えられる。(6a)の文よりも(6b)の文のほうが話者が事象に対して強い確信を持っていると言える。そこで、応用タスクに対する有用性を高くするため、確信の度合いに対応した体系の設計を試みた。確信の度合いをどのように決定するかは議論の余地があるが、本研究では暫定的に確信度の違いを元に大きく2つに分類し、確信度の高いものに「高確実性」、確信度の低いものに「不確実」の意味ラベルを割り当てた。これらの表現は、つつじでは「推量」の意味ラベルのみが割り当てられる。

また、「受身」「無意志」と意味ラベルも新たに定義した意味ラベルであるが、これらはつつじ内に対応する意味ラベルが存在しない。

(7) a. 一定時間経過後削除され受身 てしまいますよ。

b. 2点減点の違反をしてしまし無意志 た。

c. もう食べてしまっ完了 た。

(7a)の例には、助動詞「れる」が含まれる。つつじは文法形式のうちヴォイスに対応していないため、「れる」に対して一括して「可能」の意味ラベルが割り当てられる。本研究では、受動態と能動態の区別をするために、新たに「受身」ラベルを定義した。また、(7b)、(7c)の例には、「てしまう」という機能表現が含まれる。いずれも事象の完了を表す表現であるが、(7b)では事象が完了したことよりも、「主体の意志によるものではないこと」に焦点が置かれている。この情報を区別するために、「無意志」ラベルを新たに定義した。

最終的に、各カテゴリと付与対象を定義したアノテーションガイドラインを作成し、ウェブ上で公開している²。

²http://www.cl.ecei.tohoku.ac.jp/japanese_fe_corpus/fe_manual_ver2.1.pdf

3.3 機能表現意味ラベル付与コーパスの構築

我々は、現代日本語書き言葉均衡コーパス (BCCWJ)³内のテキストのうち Yahoo!知恵袋 (OC) ドメインに属するテキストを対象に、5,736 文、20,488 個の機能表現に前章で述べた意味体系に基づく意味ラベルを人手で付与した機能表現意味ラベル付与コーパスを構築した。BCCWJ は、著作権処理を実施したコーパスであり、ラベル付与結果を共有することができる。本研究では、他ジャンルに比べてウェブテキストにおいてモダリティ表現が多く、拡張モダリティ解析の恩恵が大きいと考え、BCCWJ 内の Yahoo!知恵袋 (OC) を対象に意味ラベルを付与した。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、これを形態素解析器 MeCab[22] と構文解析器 CaboCha[23] を用いて再解析した結果を用いる。

3.3.1 意味ラベルの表現方法

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 [24] 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、前章で述べた 72 種類の意味カテゴリを用いる。表 1 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「。」などの記号にはラベル 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-不確実」のように付与し、後続の形態素に「I-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、述部の機能表現ではないため、ラベル O を付与する。

³http://pj.ninjal.ac.jp/corpus_center/bccwj/, Accessed: 2015-11-10

表 1: 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-不確実
しれ	I-不確実
ない	I-不確実
。	O

3.3.2 アノテーション作業

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1名の日本語母語話者が行った。アノテーションガイドラインは付与作業を行いながら作成した。また、作業間一致率を算出するために、ガイドラインが完成した後に外部の業者に委託し、作業員2名によるアノテーションを行った。

作業員は、アノテーションガイドラインに従い、(i) 文中の述語をマークし、(ii) 述語に付属する機能表現のチャンキングを行い、ラベルBまたはIを付与し、(iii) 各機能表現に対して適切な意味ラベルを付与した。(iv) 最後に、以上の手順において機能表現であることを表すラベルB, I, または述語ラベルのいずれも付与されなかった形態素に機械的にラベルOを付与した。先述のように、複合辞には形式名詞などを含むことがあり、拡張モダリティタグ付与コーパスにおいて「事象」とされる表現と述語の認定に揺れがあるため、機能表現と同時に述語もマークした。述語の判定基準は、NAIST テキストコーパスのアノテーション仕様書における「述語」の項に準拠した。また、判断の難しい事例はその都度、備考欄にその旨を記述した。そのような事例を表2に示す。

(a) のように口語体で書かれた事例は形態素解析を誤ることが多く、形態素単

表 2: 作業者が判断に迷った事例

	例文	備考
(a)	... を見つめてる時って何を考えているのでしょうか？ ... 一番端の下の方に金額を書けばいいんでしょうか？	形態素解析誤り
(b)	でしょうか。	不完全な文
(c)	... 自ら社長室へ行き意見を聞いたりしています。 ... もう別にいっかとも思ったりもするんですが ...	複合辞性の低い事例
(d)	みんな凄く育児が楽し、上の子が、 ... こんな私は母失格ですよね。	意味を一意に決めがたい事例
(e)	皆さんのお子さんは何ヶ月位から笑い始めましたか？	機能表現の範疇に含めるかの判断が難しい例

位で意味ラベルを付与できなかった。本論文執筆時点では、こうした事例は意味ラベル付与の対象外としている。また、アノテーション対象の文には、(b)のように文の区切りが不適切と思われるものも含まれていた。機能表現としての単位が適切なものには意味ラベルを付与し、そうでない文は付与の対象外とした。(c)は複合辞の認定に関する事例である。「たりする」という形での出現例が多く、複合辞として認めることも可能である。一方で、「たりもする」のように語の挿入を許す表現もあるため、複合辞性の観点から複合辞とはせず、「たり」に例示ラベルを付与するに留めた。(d)の例は、明らかに機能表現であるが、意味ラベルの選択が難しい事例である。機能表現「し」には、並立または理由の意味ラベルが付与される可能性があり、機能表現「よね」は、態度ラベルと疑問ラベルが候補に上がる。これらは文脈によって判断が左右される事例であり、その判断基準を明確に決めがたい。本論文では、文脈に依存するラベルの間で判断に迷った場合は、付与するラベルの優先度を事例ごとに指定した。(e)のような例は、機能表現とするか接尾辞とするかの判断に迷った事例であったが、本論文では機能表現とはせず、意味ラベル付与の対象外とした。

3.4 機能表現意味ラベル付与コーパス

上記のアノテーション作業の後、意味ラベル付与結果を事実性解析器に適用し、そのフィードバックと作業中に挙げた課題をもとにアノテーション仕様を改善

するという過程を数回繰り返した。Yahoo!知恵袋ドメインのうちコアデータとされている 5,436 文を用いて仕様設計とアノテーション作業を行い、ガイドラインおよび機能表現辞書を得た。この 5,436 文を以降では訓練データとする。また、ガイドラインに基づき、コアデータ以外からランダムにサンプリングした 300 文にアノテーションを施し、これを評価データとした。

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 3 に示す。意味ラベルが付与された機能表現の延べ総数が 20,488 個、機能表現の異なり総数は、訓練データが 825 個、評価データが 183 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、訓練データが 27,190 個、評価データが 1,592 個である。訓練データ、評価データともに全形態素数に対して約 27%が本研究で対象とする機能表現を構成する形態素だった。

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、それぞれ 5,393 個、291 個であった。複合辞の占める割合は 25%~28%に上り、機能表現解析を行う上で複合辞の認識は大きな課題の 1 つになる。

表 3: 機能表現意味ラベルコーパスの統計情報

項目	訓練データ	評価データ
文数	5,436	300
形態素数	97,943	5,973
述部数	11,594	677
機能表現数 (token)	19,334	1,154
機能表現数 (type)	825	183
複合辞数	5,393	291

複合辞に関して調査するために、構成する形態素数別の複合辞数を表 4 にまとめた。例えば、「で. は. ない」という複合辞の形態素数は 3 である。形態素数が 1 の場合は複合辞の定義から外れるが、表 4 では、複合辞と合わせて記載している。結果として形態素数が多い複合辞ほど出現数は少なくなったが、もっとも形態素数の多いものでは「の. ではない. でしょうか」など、7 個の形態素からなる機能表現もみられた。形態素数別の例を (8) に示す。

- (8) a. 絶対に [ウソ] です判断。
- b. ... 一番最初にでたソフトって [何] です.か疑問 ??
- c. また最後を [見逃し] て.しまい.まし無意志 た。
- d. 他の回答者のせいで [混乱する] かも.しれ.ませ.ん不確実 が、...
- e. ... 画質が [悪く] て.も.かまい.ませ.ん許可
- f. 仕事の能力と年収が [比例する] 訳.で.も.あり.ませ.ん否定 しね。
- g. ... お花見らしい [お弁当] になる の.で.は.ない.でしょ.う.か疑問

表 4: 構成する形態素数別の複合辞数

形態素数	訓練データ	評価データ
1	13,941	862
2	3,650	188
3	1,196	67
4	427	28
5	79	7
6	29	1
7	12	0

また、述部あたりの機能表現数は、表5のようであった。述部あたりの機能表現数とは、述語の後に連続する機能表現の数を指し、複合辞はひとつと数える。すなわち、「壊し/てしまっ/た/かもしれない」では、「壊す」という述語に「てしまっ」「た」「かもしれない」の3つの機能表現が連続すると数える。連続する機能表現数別の例を(9)に示す。機能表現数が多いものでは、(9f)のようにひとつの述部に6個の機能表現を含む場合があった。

- (9) a. 知ってるかた [教え] て下さい依頼 !!

- b. ... 今年の [年末] とのこと 伝聞 です 判断。
- c. ... この中の [どれ] だっ 判断 た 完了 んです 判断 か 疑問 ?
- d. ... [つけ] た方がよかつ 勧め た 完了 です 判断 ね 態度。
- e. 今日も面接に [行こ] うとし 意志 てい 結果状態 た 完了 のです 判断 が 逆接確定...
- f. 姿勢をよくさせるために長刀を [やら] せ 使役 てい 継続 た 完了 そう 伝聞 です 判断 から 理由。

表 5: 事象に連続して付属する機能表現数

機能表現数	訓練データ	評価データ
1	5,853	300
2	4,141	237
3	1,277	93
4	261	14
5	49	3
6	12	0

3.5 作業者間一致率

作業者間の意味ラベル付与の一致率を調査するため、ガイドライン完成後に機能表現の知識のない外部の業者に委託し、新たに作業者2名によるアノテーションを行った。アノテーション対象は、訓練データから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文末の述部の機能表現のみに限定した。

一致率は、述語の一致、機能表現抽出の一致、意味ラベルの一致の3つの観点から κ 値 [25] を算出した。述語の一致は、2名の選択した述語が一致するかを評価する。機能表現抽出の一致については、述語が一致した場合に、述語に続く機

能表現の抽出の一致を評価する．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 6 に示す．

表 6: 作業員間一致率の κ 値

項目	κ
述語	0.851
機能表現抽出	0.971
意味ラベル	0.851

全体として高い一致率を達成した．一致しない事例を調査すると，想定していた意味以外で使用される場合に機能表現としてラベル付与された場合に，述語の不一致が生じやすいことが分かった．例えば「たらよい」という機能表現について，(10b) のように用いられる場合を想定し，意味カテゴリ「勧め」を定義したが，(10a) の場合にも付与される事例が多く見受けられた．NAIST テキストコーパスのアノテーション仕様書に基づく判定により述語判定は概ね一致するが，機能表現との区別についてより詳しいガイドラインを作成する必要があることが明らかになった．

(10) a. どうしたらよいのでしょうか。

b. その事だったら田中さんに相談したらよい。

一方で，機能表現抽出の一致率は非常に高く，述語判定が一致すると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，ガイドラインの記述によって改善が見込める．

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，以下の式に与えられる F 値による作業員間一致率も評価した．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7に示す．

$$Precision = \frac{\text{作業員 } B \text{ が正しく意味ラベルを付与した機能表現数}}{\text{作業員 } B \text{ が意味ラベルを付与した機能表現数}}$$

$$Recall = \frac{\text{作業員 } B \text{ は正しく意味ラベルを付与した機能表現数}}{\text{作業員 } A \text{ が意味ラベルを付与した機能表現数}}$$

$$F_{\beta=1} = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

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が存在することが分かった．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だった．これらは以下の機能表現に付与されるラベルであり，一方の作業員が付与した意味ラベルに対して，もう一名の作業員が決まって別のラベルを付与した事例である．

- (11) a. ... 防止機能の有無などが出 てきます_{方向。} / てきます_{着継続。}
- b. 探し方が悪い の_{名詞化} かな_{疑問？} / の_{かな} 疑問？
- c. ... どこかのページにのっているなら教え ていただき_{受益} / ていただき_{依頼} たいです_{願望。}
- d. 使わない方がいいか と_{内容。} / と_{判断。}
- e. ... 対応して入る 方_{比較} がいいです_{願望。} / 方がいいです_{勧め。}
- f. ... 原作通りで なくてもいい_{不必要} / なくてもいい_{ラベルなし} ものなのですか？

その他に低い一致率を示すのは，「結果状態」「習慣」「継続」といった意味ラベルである．これらは，(12)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，表層を同じとする表現の曖昧性による不一致である．現在の仕様では，これらの間で判断に迷った場合は，定められた優先度にしたがって意味ラベルを選択するという方針をとっている．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

することが難しい。意味ラベルの粒度の見直しか，言語現象のより詳しい分析が必要である。

(12) 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 7: 意味ラベルの作業者間一致率

意味ラベル	Precision	Recall	$F_{\beta=1}$	意味ラベル	Precision	Recall	$F_{\beta=1}$
疑問	92.81 (297/320)	94.59 (297/314)	93.69	許可	100.0 (5/5)	83.33 (5/6)	90.91
判断	92.86 (247/266)	95.37 (247/259)	94.10	試行	85.71 (6/7)	100.0 (6/6)	92.31
完了	80.85 (114/141)	93.44 (114/122)	86.69	可能	62.50 (5/8)	100.0 (5/5)	76.92
結果状態	54.60 (89/163)	74.79 (89/119)	63.12	伝聞	100.0 (5/5)	100.0 (5/5)	100.0
習慣	89.47 (34/38)	40.00 (34/85)	55.28	当為	50.00 (3/6)	75.00 (3/4)	60.00
態度	90.79 (69/76)	88.46 (69/78)	89.61	様態	100.0 (4/4)	100.0 (4/4)	100.0
否定	76.47 (52/68)	70.27 (52/74)	73.24	程度	60.00 (3/5)	100.0 (3/3)	75.00
受身	100.0 (39/39)	92.86 (39/42)	96.30	強調	100.0 (3/3)	100.0 (3/3)	100.0
継続	71.43 (10/14)	25.64 (10/39)	37.74	自発	40.00 (2/5)	100.0 (2/2)	57.14
話題	100.0 (38/38)	97.44 (38/39)	98.70	名詞化	0.00 (0/2)	0.00 (0/2)	0.00
無意志	82.93 (34/41)	100.0 (34/34)	90.67	受益	0.00 (0/0)	0.00 (0/2)	0.00
勧め	76.92 (10/13)	34.48 (10/29)	47.62	意志	66.67 (2/3)	100.0 (2/2)	80.00
理由	100.0 (21/21)	91.30 (21/23)	95.45	内容	0.00 (0/1)	0.00 (0/2)	0.00
願望	100.0 (12/12)	85.71 (12/14)	92.31	目的	100.0 (1/1)	50.00 (1/2)	66.67
自然発生	21.15 (11/52)	78.57 (11/14)	33.33	例示	100.0 (1/1)	100.0 (1/1)	100.0
依頼	73.33 (11/15)	100.0 (11/11)	84.62	比較	0.00 (0/0)	0.00 (0/1)	0.00
方向	0.00 (0/0)	0.00 (0/11)	0.00	不必要	0.00 (0/0)	0.00 (0/1)	0.00
不確実	100.0 (10/10)	100.0 (10/10)	100.0	容易	100.0 (1/1)	100.0 (1/1)	100.0
着継続	47.37 (9/19)	100.0 (9/9)	64.29	All	81.77 (1148/1404)	83.31 (1148/1378)	82.53

4 機能表現解析

本節では、構築した機能表現意味ラベル付与コーパスに基づく統計的学習により、どれほどの機能表現解析が実現できるかを検証する。以下では、実験設定とその結果を報告する。

4.1 条件付き確率場

本論文では、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)[26] による学習を行う。本論文における CRF は、linear-chain CRF を指し、一次マルコフ性を仮定する。CRF は識別モデルであり、以下で式で与えられる入力系列 x に対する出力系列 y の条件付き確率 $P(y|x)$ において、正しい系列に対する条件付き確率がもっとも大きくなるように構造学習を行う。

$$P(y|x) = \frac{\exp(\mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y}))}{\sum_{\mathbf{y}} \exp(\mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y}))} \quad (1)$$

ここで、 ϕ は素性関数、 \mathbf{w} は素性関数に対する重みベクトルである。素性関数 ϕ は、入力 x とラベル y を引数にとって、素性ベクトルを返す関数である。素性ベクトル中の各素性値は、以下のように定義され、現在の位置を t として、直前のラベル y_{t-1} 、現在のラベル y_t 、観測された素性 x が特定の組み合わせのときに限り 1 となる。

$$\phi_k(\mathbf{x}, y_t, y_{t-1}) = \begin{cases} 1 & (w = \text{かも} \ \& \ p = \text{助詞}) \\ 0 & (\text{otherwise}) \end{cases} \quad (2)$$

学習は、以下のように $P(y|x)$ の対数尤度を目的関数 $L(\mathbf{w})$ とした最大化問題を解き、パラメータ \mathbf{w} を推定する。

$$L(\mathbf{w}) = \sum \log P(y|x) - r(\mathbf{w}) \quad (3)$$

$r(\mathbf{w})$ は正則化項である。パラメータ推定は、準ニュートン法などが用いられることが多い。学習したパラメータを元に、入力 x に対する最適な出力 \hat{y} を以下

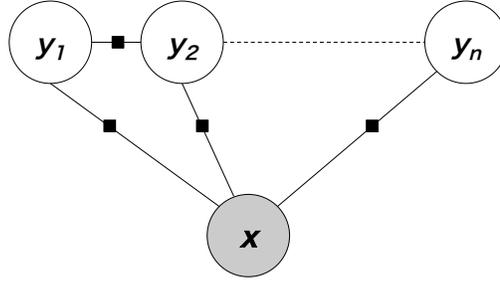


図 1: Graphical model of linear-chain CRF.

のように得ることができる．

$$\hat{y} = \arg \max_y P(\mathbf{y}|\mathbf{x}) \quad (4)$$

CRF は，固有表現認識 (Named Entity Recognition, NER) をはじめとする系列ラベリング問題に対して広く使われる手法であり，柔軟な素性設計が可能であるなどの特徴を持つ．機能表現解析においては，入力形態素列，出力意味ラベル系列となる．

4.2 評価実験

本論文では，述部の機能表現のみを解析対象としているため，入力には機能表現を構成する形態素列を与える．実験は，訓練データを用いた十分割交差検定と，訓練データを教師として評価データをテストするオープンテストの 2 種類の実験を行う．CRF の実装には CRFSuite[27] を用いた．各パラメータはデフォルト設定を用いる．すなわち，学習時には，Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) および L2 正則化を使用する．

4.2.1 学習素性

学習素性には，主に，MeCab による形態素解析で得られる単語素性を用いた．実験に仕様した素性を表 8 に示す．表の w , p , $p1$, $p2$, $p3$, cf , bf は， x_t 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， w' , p' ,

cf', bf' は, x_{t-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もって区別した.

表 8: 学習素性の一覧.

	テンプレート	例
uni-gram	$\langle w \rangle$	しれ
	$\langle p \rangle$	動詞
	$\langle cf \rangle$	未然形
	$\langle bf \rangle$	しれる
	$\langle p, p1 \rangle$	動詞 自立
	$\langle p, p1, p2 \rangle$	動詞 自立 *
	$\langle p, p1, p2, p3 \rangle$	動詞 自立 * *
	$\langle bf, p \rangle$	しれる 動詞
	$\langle bf, p, p1 \rangle$	しれる 動詞 自立
bi-gram	$\langle bf, p, p1, p2 \rangle$	しれる 動詞 自立 *
	$\langle w', w \rangle$	かも しれ
	$\langle p', p \rangle$	助詞 動詞
	$\langle bf', bf \rangle$	かも しれる
	$\langle cf', w \rangle$	* かも

4.2.2 Baseline: 最長一致法

ベースラインとして, 辞書を用いた最長一致法による付与結果を用いた. 機能表現辞書は, 表層表現と対応する意味ラベルを収録したもので, 訓練データに含まれる機能表現をもとに構築したものを使用する. ベースラインシステムは, 入力文の前方の形態素から機能表現辞書を検索し, 対象形態素で始まる機能表現のうち, もっとも形態素数が多い機能表現を選び, 対応する意味ラベルを付与する. 対応する意味ラベルが複数ある場合には, 候補の意味ラベルのうち, 訓練データ内でもっとも出現頻度が高い意味ラベルと選択する. 実験では, 辞書検索の際に直前の形態素の単語素性 (品詞, 品詞細分類, 活用型, 活用形, 基本形, 読み, 発音) を制約を考える. 制約なしのシステムを Baseline A, 制約を用いて素性が一致しない場合に候補から除外するシステムを Baseline B とする. 制約情報は訓練データを元に作成したものをを用いる.

4.3 評価

4.3.1 評価尺度

評価は、以下で与えられる Precision, Recall, および, F 値によって行う。なお, 評価単位は, 機能表現, すなわち, BI ラベルの連続を一単位とする。正解の基準として, 機能表現区切りのみの正解, 機能表現区切りと意味ラベルが正解の 2 つを設けた。機能表現区切りは, 機能表現の範囲を表す B または I のラベルが一致していれば正解とする。

$$\begin{aligned} Precision &= \frac{\text{正しく意味ラベルが付与された機能表現数}}{\text{いずれかの意味ラベルが付与された機能表現数}} \\ Recall &= \frac{\text{正しく意味ラベルが付与された機能表現数}}{\text{評価データに存在する機能表現数}} \\ F_{\beta=1} &= \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \end{aligned}$$

4.3.2 評価結果

機能表現解析器の評価結果を表 9, 表 10 に示す。全体として, CRF モデルによる解析性能は, クローズドテストとオープンテストのいずれの場合にも 83% 以上の F 値を達成しており, 構築した意味ラベル付与コーパスは学習できるだけの一貫性があると言える。

機能表現区切りのみの評価では, 接続制約を用いた Baseline B によるオープンテストで 9 割を下回ったものの, それ以外の条件では 94% と高い F 値を達成した。意味ラベルの一致を見ると, クローズドテストでは Baseline B が最も良く, CRF モデルが Baseline B をやや下回る結果であった。一方, オープンテストでは, Baseline B の精度は大きく下がり, CRF モデルが Baseline B の精度を上回った。また, 最長一致法を用いた場合, クローズドテストでは接続制約が解析性能に大きく貢献したが, オープンテストでは接続制約を用いない方が解析性能が高い。これは接続制約が厳しく, 未知の機能表現を正しく認識できていないことが問題であると考えられる。接続制約の表現方法にも改善の余地はあるが, CRF モデルの単語素性が接続制約に相当すると考えれば, より汎化能力の高い CRF モデルのほうが機能表現解析に適していると言える。

表 9: 機能表現解析器の評価結果 (Closed)

	Method	Precision		Recall		$F_{\beta=1}$
機能表現区切り	Baseline A	99.89	(19103/19124)	98.81	(19103/19334)	99.34
	Baseline B	99.38	(17939/18051)	92.78	(17939/19334)	95.97
	CRF	96.71	(18806/19446)	97.27	(18806/19334)	96.99
意味ラベル	Baseline A	74.22	(13398/18051)	69.30	(13398/19334)	71.68
	Baseline B	84.71	(16200/19124)	83.79	(16200/19334)	84.25
	CRF	83.20	(16179/19446)	83.68	(16179/19334)	83.44

表 10: 機能表現解析器の評価結果 (Open)

	Method	Precision		Recall		$F_{\beta=1}$
機能表現区切り	Baseline A	97.58	(1050/ 1076)	90.99	(1050/ 1154)	94.17
	Baseline B	91.10	(911/ 1000)	78.94	(911/ 1154)	84.59
	CRF	98.95	(1136/1148)	98.44	(1136/1154)	98.70
意味ラベル	Baseline A	74.54	(802/ 1076)	69.50	(802/ 1154)	71.93
	Baseline B	75.20	(752/ 1000)	65.16	(752/ 1154)	69.82
	CRF	88.33	(1014/1148)	87.87	(1014/1154)	88.10

4.4 議論

より詳細な分析のために、表 11 に意味ラベル別の性能を示す。十分な数の事例がある意味ラベルは比較的高い精度で解析できた。事例数が多いにも関わらず精度が伸び悩んだ意味ラベルとして、「結果状態」がある。この意味ラベルは、作業員間一致率の項で述べたように人手のアノテーションも一致率が低いラベルであった。すなわち、意味体系の設計に改善の余地があり、現象と照らしあわせて粒度の再検討、手がかりの分析を行う必要がある。「順接確定」ラベルも同様な理由で精度が低い。こちらは主に、「て」という機能表現に付与される意味ラベルである。また、「最中」や「経歴」など出現頻度の低い意味ラベルは事例数が少なく、十分に学習することができなかった。

本論文で粒度の再考を行った意味ラベルとして「高確実」と「不確実」があった。いずれのラベルも精度には改善の余地が残った。誤りの原因を調査するため、特に「高確実」ラベルについて、システムが機能表現を正しく認識できなかった事例を分析し、これを二つに分類した。ひとつ目は、正解データでは付与されてい

ない意味ラベルを誤って付与した事例，ふたつ目は，正解データでは意味ラベルが付与されているにも関わらず正しい意味ラベルを付与できなかった事例である．

表 11: 意味ラベル別の評価結果 .

意味ラベル	Precision	Recall	$F_{\beta=1}$	意味ラベル	Precision	Recall	$F_{\beta=1}$
判断	93.55 (4266/4560)	95.50 (4266/4467)	94.52	許可	72.88 (43/59)	68.25 (43/63)	70.49
完了	91.26 (1451/1590)	95.59 (1451/1518)	93.37	比較	81.82 (27/33)	58.70 (27/46)	68.35
疑問	92.27 (1396/1513)	93.00 (1396/1501)	92.63	比況	52.94 (27/51)	58.70 (27/46)	55.67
否定	90.76 (1110/1223)	96.10 (1110/1155)	93.36	使役	82.05 (32/39)	72.73 (32/44)	77.11
理由	84.89 (753/887)	85.86 (753/877)	85.37	着継続	60.00 (24/40)	54.55 (24/44)	57.14
結果状態	62.21 (507/815)	66.02 (507/768)	64.06	容易	90.24 (37/41)	86.05 (37/43)	88.10
内容	91.12 (739/811)	96.35 (739/767)	93.66	付帯-並行	82.76 (24/29)	57.14 (24/42)	67.61
態度	91.14 (669/734)	91.64 (669/730)	91.39	様態	33.33 (8/24)	21.05 (8/38)	25.81
順接仮定	76.54 (535/699)	82.31 (535/650)	79.32	当為	91.30 (21/23)	61.76 (21/34)	73.68
逆接確定	66.49 (379/570)	63.91 (379/593)	65.18	授与	95.65 (22/23)	75.86 (22/29)	84.62
依頼	90.50 (505/558)	91.99 (505/549)	91.24	受益	57.89 (11/19)	47.83 (11/23)	52.38
順接確定	43.44 (225/518)	49.56 (225/454)	46.30	強調	50.00 (4/8)	19.05 (4/21)	27.59
受身	86.73 (438/505)	96.48 (438/454)	91.35	困難	100.00 (13/13)	72.22 (13/18)	83.87
名詞化	90.79 (414/456)	95.83 (414/432)	93.24	感嘆	66.67 (4/6)	26.67 (4/15)	38.10
話題	61.36 (262/427)	64.22 (262/408)	62.75	反復	0.00 (0/0)	0.00 (0/15)	0.00
継続	56.92 (185/325)	48.43 (185/382)	52.33	不可避	46.67 (7/15)	46.67 (7/15)	46.67
自然発生	75.60 (282/373)	80.57 (282/350)	78.01	発継続	69.23 (9/13)	60.00 (9/15)	64.29
習慣	50.68 (112/221)	40.43 (112/277)	44.98	終点	93.33 (14/15)	100.00 (14/14)	96.55
勤め	71.73 (203/283)	73.82 (203/275)	72.76	勧誘	66.67 (2/3)	16.67 (2/12)	26.67
願望	95.02 (191/201)	88.84 (191/215)	91.83	不可能	0.00 (0/3)	0.00 (0/12)	0.00
並立	36.69 (51/139)	24.06 (51/212)	29.06	手段	66.67 (2/3)	16.67 (2/12)	26.67
例示	69.66 (101/145)	65.58 (101/154)	67.56	不許可	33.33 (2/6)	20.00 (2/10)	25.00
無意志	85.71 (132/154)	89.80 (132/147)	87.71	自発	80.00 (4/5)	44.44 (4/9)	57.14
不確実	73.51 (111/151)	76.03 (111/146)	74.75	不必要	0.00 (0/1)	0.00 (0/9)	0.00
高確実	84.50 (109/129)	80.15 (109/136)	82.26	事前	100.00 (1/1)	14.29 (1/7)	25.00
継起	61.25 (49/80)	38.58 (49/127)	47.34	事後	0.00 (0/2)	0.00 (0/5)	0.00
逆接仮定	69.86 (102/146)	80.95 (102/126)	75.00	場合	0.00 (0/0)	0.00 (0/3)	0.00
程度	87.29 (103/118)	88.03 (103/117)	87.66	起点	0.00 (0/0)	0.00 (0/2)	0.00
意志	88.00 (88/100)	86.27 (88/102)	87.13	対比	0.00 (0/0)	0.00 (0/2)	0.00
添削	75.29 (64/85)	68.82 (64/93)	71.91	同時性	0.00 (0/0)	0.00 (0/2)	0.00
付加	49.25 (33/67)	36.26 (33/91)	41.77	無意味	0.00 (0/0)	0.00 (0/2)	0.00
目的	70.87 (73/103)	80.22 (73/91)	75.26	順接限定	0.00 (0/0)	0.00 (0/1)	0.00
方向	73.03 (65/89)	72.22 (65/90)	72.63	最中	0.00 (0/0)	0.00 (0/1)	0.00
試行	95.00 (76/80)	92.68 (76/82)	93.83	経歴	0.00 (0/0)	0.00 (0/1)	0.00
可能	56.06 (37/66)	48.68 (37/76)	52.11	付帯-続行	0.00 (0/0)	0.00 (0/1)	0.00
付帯-続行	50.00 (25/50)	37.88 (25/66)	43.10				

表 12: 誤り分析

Semantic labels	gold	tp	fp	fn	P	R	F
判断	4472	4272	297	200	93.51	95.53	94.51
高確実性	136	107	17	29	86.29	78.68	82.31

「判断」ラベルを付与することが出来なかった事例は 4472 個中 200 個 (4.5%) , 誤って「判断」ラベルを付与してしまった事例は 4472 個中 297 個 (6.6%) あった . 「判断」ラベルは , ほぼ無標のラベルとして付与しているため出現数が極めて多い . したがって , 原形に「だ」「です」「ます」などの判断ラベルを付与されることの多い語を含むと , 誤って判断ラベルを付与するという事例が目立った . これを改善するためには , 判断ラベルの特徴を捉える素性を組み込むよりも , 他のラベルを正しく抽出するほうが良い策であると考えられる .

「高確実性」ラベルを付与することが出来なかった事例は , 136 個中 29 個 (21.3%) あった . これらの多くは , 「疑問」「判断」ラベルを付与されることが多かった . 例えば , (13a) の下線部の機能表現に対する意味ラベルは「高確実性」であるが , システムは誤って「疑問」ラベルを付与した . 「疑問」と「高確実性」の曖昧性は , 人間にとっても判断が難しい場合がある . これを解決するためには , より詳細な仕様が必要である . また , (13b) の下線部の機能表現には , 誤って「判断」のラベルを付与した . これは , データ中の「原形が「です」」という表現の多くに対して「判断」ラベルが付与されることが原因であると考えられる .

- (13) a. 生活の知恵な んだらう ね。 (正解:高確実性, システム:疑問)
 b. 去年の終わりから でしょ。 (正解:高確実性, システム:判断)
 c. どういう基準で外す んでしょう。 (正解:疑問, システム:高確実性)
 d. 求めない方がいい でしょう ね。 (正解:判断, システム:高確実性)

表 13 には複合辞の長さごとの評価結果を示す . 複合辞の長さは , 表 4 と同様に複合辞を構成する形態素数を指す .

表 13: 機能表現の長さ別の評価結果

機能表現長	Precision	Recall	$F_{\beta=1}$
1	84.89 (11926/14049)	85.57 (11926/13937)	85.23
2	79.84 (2942/3685)	80.51 (2942/3654)	80.17
3	71.65 (872/1217)	72.91 (872/1196)	72.28
4	90.15 (357/396)	83.61 (357/427)	86.76
5	83.61 (51/61)	64.56 (51/79)	72.86
6	76.67 (23/30)	79.31 (23/29)	77.97
7	100.00 (8/8)	66.67 (8/12)	80.00

5 拡張モダリティ解析

機能表現を用いて拡張モダリティ解析を行うシステムに本研究の機能表現解析結果を適用することで、拡張モダリティ解析における機能表現解析の寄与を検証する。本節では、まず、拡張モダリティタグと拡張モダリティタグ付与コーパスについて記述し、機能表現を用いた拡張モダリティ解析の評価実験について述べる。

5.1 拡張モダリティタグ

拡張モダリティは、松吉らによって定義された書き手の態度や真偽判断などに関する事象に対する判断情報である。ここで、事象は行為、出来事、状態の総称である。例えば、以下の例文では、「雨が降るコト」が事象であり、事象の真偽判断は「成立」、書き手の態度は「叙述」である。

(14) a. 昨日は雨が [降っ] た。

松吉らは、態度表明者、時制、仮想、態度、真偽判断、価値判断、焦点の7項目に対するタグを拡張モダリティタグとして定義した。各項目に対する拡張モダリティタグの一覧を表14に示す。また、松吉らの研究では、人手で拡張モダリティタグを付与したコーパスを構築している。コーパス内のタグの分布を表14に示す。

表 14: 拡張モダリティタグの一覧

項目	タグ	訓練データ		評価データ	
時制	非未来	10811	(3790)	898	(239)
	未来	2050	(754)	155	(52)
	0	5	(0)	0	(0)
態度	許可	24	(6)	10	(4)
	欲求	237	(48)	22	(2)
	意志	320	(51)	14	(3)
	問いかけ	1335	(968)	75	(50)
	叙述	10023	(2948)	841	(191)
	働きかけ-間接	448	(177)	37	(11)
	働きかけ-直接	462	(338)	46	(26)
	働きかけ-勧誘	12	(8)	8	(4)
真偽判断	0	5	(0)	0	(0)
	高確率から低確率	3	(0)	0	(0)
	高確率	916	(241)	26	(8)
	成立から不成立	25	(6)	2	(2)
	成立	8026	(2359)	713	(155)
	低確率から高確率	13	(2)	13	(1)
	低確率	154	(48)	18	(5)
	不成立から成立	65	(20)	3	(2)
不成立	852	(305)	75	(20)	
価値判断	0	2812	(1563)	203	(98)
	ポジティブ	1410	(594)	128	(46)
	ネガティブ	252	(80)	9	(4)
	0	11204	(3870)	914	(241)

5.1.1 ロジスティック回帰による拡張モダリティ解析

本論文では，ロジスティック回帰モデルに基づく点推定により，各項目に対して多クラス分類を行う．すなわち，項目ごとに，あるタグが付与される確率 p_i を以下の式によって推定し，もっとも確率の高いタグを出力に選ぶ．ここで， x は素性， β は素性に対する重みである．

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} \quad (5)$$

5.2 実験

学習データに本研究で構築した機能表現意味ラベル付与コーパスを用いて，主節に含まれる事象（以下，主事象）の拡張モダリティ解析を行う．実験では，文末にもっとも近い述語を主事象とした．対象を主事象に限定するのは，主事象の拡張モダリティが概ね機能表現によって決定可能であるという仮説に基づく．意味ラベル付与コーパスは，松吉らによる拡張モダリティタグ付与コーパスと同じデータを対象としているが，事象の認定に不一致のある文を含むため，主事象が一致している 4,429 事象のみを解析対象とする．

拡張モダリティ解析ツールとして，日本語拡張モダリティ解析器 Zunda[28] を使用する．Zunda における学習素性を表 15 に示す．

表 15: 拡張モダリティ解析の学習素性

素性	説明
基本素性	対象形態素の前後 3 単語までの表層, および基本形
機能語列	対象形態素から文末までの形態素の n-gram
意味ラベル	対象形態素から文末までの機能表現に付与された機能表現意味ラベルの n-gram

表 16: 4 項組の評価結果

学習	解析時	素性	4 項組正解率
正解	正解	基本素性+機能語列+意味ラベル	68.73 (200/291)
正解	CRF	基本素性+機能語列+意味ラベル	68.73 (200/291)
正解	正解	基本素性+機能語列	68.04 (198/291)
正解	CRF	基本素性+意味ラベル	64.95 (189/291)
正解	正解	基本素性+意味ラベル	64.60 (188/291)
正解	正解	基本素性	62.54 (182/291)
正解	正解	基本素性+機能語列+つつじラベル	68.04 (198/291)
正解	正解	基本素性+つつじラベル	63.23 (184/291)
交差検定	CRF	基本素性+機能語列+意味ラベル	68.04 (198/291)
交差検定	正解	基本素性+機能語列+意味ラベル	67.35 (196/291)
交差検定	CRF	基本素性+意味ラベル	65.29 (190/291)
交差検定	正解	基本素性+意味ラベル	63.92 (186/291)

学習および解析時の機能表現データ, 学習素性別の正解率。(上) 本研究における意味ラベルを用いた場合, (中) つつじラベルを用いた場合, (下) 交差検定結果を学習した場合. 太字は各段における上位 1 位.

実験は, i) 学習素性の比較, ii) 本研究の意味体系を用いた場合と既存辞書の意味体系を用いた場合の比較, iii) 正解ラベルを用いた場合と機能表現の自動解析結果を用いた場合の比較の 3 項目について行う.

拡張モダリティ解析では, 訓練データを用いた十分割交差検定を実施し, 以下で与えられる Precision, Recall, F 値によって評価する.

$$\begin{aligned}
 Precision &= \frac{\text{正しく拡張モダリティタグが付与された事象数}}{\text{拡張モダリティタグが付与された事象数}} \\
 Recall &= \frac{\text{正しく拡張モダリティタグが付与された事象数}}{\text{評価データに存在する事象数}} \\
 F_{\beta=1} &= \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}
 \end{aligned}$$

5.3 評価結果

学習素性を変えた場合，機能表現の意味体系を変えた場合，Zunda の学習および解析時に使用する機能表現意味ラベルを変えた場合の時制，態度，真偽判断，価値判断の 4 項組について，全ての項目が正解したものを正解とみなしたときの正解率を表 16 に示す．評価データの 300 文のうち，9 文には事象が含まれなかったため，291 事象が評価対象である．

5.3.1 学習素性の比較

学習および解析時に使用する意味ラベルとして人手による正解データを用い，学習素性を変えた結果を比較すると(上段)，機能語列が大きく貢献しており，次いで意味ラベルが有効であることが確かめられた．基本素性に加えてこれら 2 つの素性を組み合わせることによってさらに正解率は高くなった．機能語列の情報だけでなく，機能表現を意味の単位で抽象化した情報を用いることによって拡張モダリティ解析の精度向上を確認できた．

5.3.2 意味体系の比較

機能表現の意味体系を変えた場合を比較すると(上段および中段)，基本素性と意味ラベルを組み合わせただけの場合にはいずれの意味体系を用いた場合にも精度の改善は見られたが，さらに機能語列を加えた場合には既存の意味体系では精度向上が見られなかった．

5.3.3 意味ラベルの付与方法の比較

Zunda の学習時に使用する機能表現データとしては，交差検定結果を用いるよりも正解データを用いるほうが高い精度で解析できた(上段および下段)．学習に正解データ，学習素性として全素性を用いた場合には，解析時の機能表現データは CRF の解析結果を用いた場合でも正解データを用いた場合と同じ正解率を達成した．機能語文字列素性を用いない場合や学習データとして交差検定結果を用

いる場合には CRF の解析結果のほうが正解率が高かった。これは出現率の低い機能表現に対して Zunda が学習できなかった情報を CRF が補ったと考えられる。拡張モダリティ解析を行う場合には、機能表現解析は現実的な解析性能を達成していると言える。

5.3.4 項目ごとの評価結果

機能表現データとして正解データを用いた場合について、各項目ごとに最頻出のタグを除くマクロ F 値を表 17 に示す。最頻出のタグを除くのは、最頻出のタグが大部分を占めているが、拡張モダリティ解析の目的は最頻出のタグ以外にも正しく解析することであるためである。

表 17: 最頻出のタグを除く項目ごとの評価結果

素性	時制	態度	真偽判断	価値判断
基本素性	0.442	0.530	0.669	0.420
基本素性+意味ラベル	0.500	0.610	0.706	0.460
基本素性+つつじラベル	0.404	0.520	0.647	0.460
基本素性+機能語列	0.462	0.620	0.706	0.460
全素性	0.442	0.620	0.713	0.460
全素性(つつじ)	0.462	0.620	0.691	0.460

太字は項目ごとの上位 1 位。学習および解析時の意味ラベルは正解データを用いた。

時制は意味ラベルのみを学習素性としたとき、もっとも高い精度で解析でき、機能語列を加えると精度が低下した。真偽判断は、意味ラベルと機能語列が同程度の貢献を示しており、これらを組み合わせることによってさらに高い精度を達成した。態度や価値判断は意味ラベルの貢献はあるものの、機能語列のみを用いた場合を大きく上回ることはなかった。この結果は、真偽判断を中心に意味体系の設計を行ったことから妥当であると言える。同時に、項目ごとに意味ラベル素性の貢献が異なることから項目別に素性設計を行うことも有効であることが伺える結果であった。

6 おわりに

本論文では、機能表現の意味ラベル体系を新たに設計し、作成した体系に基づき日本語コーパス中の機能表現に対して意味ラベルを人手で付与した。構築したコーパスは、5,736 文、20,488 個の機能表現を含み、作業者間一致率は $F=83.98\%$ と十分に高い値を達成した。アノテーションガイドラインおよび構築したコーパスはウェブ上で公開し、一般に利用可能である。

CRF による機能表現解析では、 $F=88.10\%$ を達成し、作成したコーパスと機械学習を用いて現実的に解析が可能であることを示した。非常に単純な学習素性のみを用いたため、機能表現解析器には大いに改善の余地がある。

さらに、本研究の機能表現解析結果を拡張モダリティ解析に適用した。日本語拡張モダリティ解析器 `zunda` の学習素性として本研究の機能表現解析結果を導入することによって性能向上が認められ、機能表現解析が拡張モダリティ解析に有効であることを示した。

謝辞

本研究を進めるにあたり，多くの方々にご協力をいただきました．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と岡崎直観准教授には，お忙しい中，研究活動全般にわたり温かいご指導，ご助言をいただきました．研究内容だけではなく，研究者としての考え方もご指導いただき，自分の価値観が大きく変わる研究生活を過ごすことができました．深く感謝致します．

水野淳太研究員には，日常的に研究の相談にのっていただくとともに，論文の添削から研究発表の指導においても大変お世話になりました．また，研究以外でも多く場面で暖かく支えてくださいました．本当に感謝致します．

研究に関して貴重なアドバイスをくださり，研究生活を暖かく支えてくださった研究室の皆様，そして大学生活において貴重かつ有意義な時間を共に過ごしてくださった皆様に心より感謝いたします．先輩方には研究うに関して議論させて頂き，またそれ以外にも学生生活に関する多くの助言を頂きました．優秀な後輩たちには多くの刺激をもらいました．同期とは多くの苦労を共にしました．皆様，本当にありがとうございました．

ご多忙の中，審査委員をお引き受けくださいました田中和之教授，伊藤彰則教授に深く感謝致します．

最後に，私の研究生活を様々な面でさせてくれた数多くの先輩，友人，知人，そして家族に心より感謝致します．

参考文献

- [1] 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 回年次大会予稿集, pp. 2–5, 2011.
- [2] 鈴木敬文, 阿部佑亮, 宇津呂武仁, 松吉俊, 土屋雅稔. 代表・派生関係を利用した日本語機能表現の解析方式の評価. 言語処理学会第 18 回年次大会予稿集, pp. 598–601, 2012.
- [3] 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.
- [4] Kazuya Narita, Junta Mizuno, and Kentaro Inui. A lexicon-based investigation of research issues in Japanese factuality analysis. In *In Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP 2013)*, pp. 587–595, 2013.
- [5] 土屋雅稔, 宇津呂武仁, 松吉俊, 佐藤理史, 中川聖一. 日本語複合辞用例データベースの作成と分析. 情報処理学会論文誌, Vol. 47, No. 6, pp. 1728–1741, 2006.
- [6] 国立国語研究所. 現代語複合辞用例集. 2001.
- [7] György Szarvas, Veronika Vincze, Richárd Farkas, and János 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.
- [8] 土屋雅稔, 注連隆夫, 高木俊宏, 内元清貴, 松吉俊, 宇津呂武仁, 佐藤理史, 中川聖一. 機械学習を用いた日本語機能表現のチャンキング. 自然言語処理, Vol. 14, No. 1, pp. 111–138, 2007.
- [9] 長坂泰治, 宇津呂武仁, 松吉俊, 土屋雅稔. 大規模階層辞書を利用した日本語機能表現の集約と解析. 言語処理学会第 15 回年次大会論文集, pp. 328–331, 2009.

- [10] Takafumi Suzuki, Yusuke Abe, Itsuki Toyota, Takehito Utsuro, Suguru Matsuyoshi, and Masatoshi Tsuchiya. Detecting Japanese Compound Functional Expressions using Canonical/Derivational Relation. In *Proceedings of the 8th International Language Resources and Evaluation*, 2012.
- [11] 注連隆夫, 土屋雅稔, 松吉俊, 宇津呂武仁, 佐藤理史. 日本語機能表現の自動検出と統計的係り受け解析への応用. *自然言語処理*, Vol. 14, No. 5, pp. 167–197, 2007.
- [12] 坂本朋子, 宇津呂武仁, 松吉俊. 日本語機能表現の集約的英訳. *言語処理学会第15回年次大会論文集*, pp. 654–657, 2009.
- [13] 島内蘭, 長坂泰治, 坂本朋子, 宇津呂武仁, 松吉俊. 日英特許翻訳における日本語機能表現の集約的英訳可能性の調査. *言語処理学会第16回年次大会論文集*, pp. 611–614, 2010.
- [14] 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.
- [15] Sanda Harabagiu, Andrew Hickl, and Finley Lacatusu. Negation, contrast and contradiction in text processing. In *Proceedings of the 21st national conference on Artificial intelligence*, pp. 755–762, 2006.
- [16] Kathrin Baker, Michael Bloodgood, Bonnie Dorr, Nathaniel W. Filardo, Lori Levin, and Christine Piatko. A modality lexicon and its use in automatic tagging. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, pp. 1402–1407, 2010.
- [17] 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回年次大会予稿集*, pp. 147–150, 2011.

- [18] 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.
- [19] 永野賢. 表現文法の問題-複合辞の認定について-, pp. 95–120. 三省堂, 1953.
「永野賢 (1970). 伝達論にもとづく日本語文法の研究. 東京堂出版」に再録.
- [20] 松木正恵. 複合辞の認定基準・尺度設定の試み. 早稲田大学日本語研究教育センター紀要, 第2巻, pp. 27–52, 1990.
- [21] 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.
- [22] 工藤拓, 山本薫, 松本裕治. Conditional random fields を用いた日本語形態素解析. 情報処理学会自然言語処理研究会 SIGNL-161, pp. 89–96, 2004.
- [23] 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, Vol. 43, No. 6, pp. 1834–1842, 2002.
- [24] Erik F. Tjong Kim Sang. Noun phrase recognition by system combination. In *Proceedings of the Language technology Joint Conference ANLP-NAACL2000*, pp. 50–55, 2000.
- [25] Jacob Cohen. A coefficient for agreement for nominal scales. In *Education and Psychological Measurement*, Vol. 20, pp. 37–46, 1960.
- [26] John Lafferty, Andrew K. McCallum, and Fernando Pereira. Conditional random fields: probabilistic models for segmenting and labeling sequence data. In *ICML*, pp. 282–289, 2001.
- [27] Naoaki Okazaki. CRFsuite: a fast implementation of conditional random fields (CRFs), 2007.
- [28] Junta Mizuno. Zunda, 2013.

発表文献一覧

受賞一覧

- 情報処理学会第 77 回全国大会学生奨励賞 (2015)
- 情報処理学会第 77 回全国大会大会奨励賞 (2015)

学術論文誌

- 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.392-432, December 2015.

国際会議論文

- 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.

国内会議・研究会論文

- 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.
- 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.

- 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.
- 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.