

名詞句の語彙統語パターンを用いた事態性名詞の項構造解析

小町 守[†]・飯田 龍^{††}・乾 健太郎[†]・松本 裕治[†]

形態素解析や構文解析など自然言語処理の要素技術は成熟しつつあり、意味解析・談話解析といった、より高次の言語処理の研究が盛んになってきた。特に文の意味理解のためには「誰が」「何を」「誰に」といった要素（項）を同定することが重要である。動詞や形容詞を対象にした項構造解析のことを述語項構造解析と呼ぶが、文中の事態を表す表現は動詞や形容詞の他にも名詞も存在することが知られている。そこで、我々は日本語の名詞を対象とした項構造解析タスクを取り上げ、機械学習を用いた自動的な解析手法を提案する。日本語の事態性名詞には事態を指すか否か曖昧性のある名詞があるため、まず事態性の有無を判定する事態性判別タスクと項同定タスクの2つに分解し、それぞれ大規模なコーパスから抽出した語彙統語パターンを用いた手法と述語・事態性名詞間の項の共有現象に着目した手法を提案する。

キーワード：述語項構造解析, 意味役割付与, 事態性名詞, 語彙統語パターン

Argument Structure Analysis of Event-nouns Using Lexico-syntactic Patterns of Noun Phrases

MAMORU KOMACHI[†], RYU IIDA^{††}, KENTARO INUI[†] and YUJI MATSUMOTO[†]

As fundamental natural language processing techniques like morphological analysis and parsing have become widely used, semantics and discourse analysis has gained increasing attention. Especially, it is essential to identify fundamental elements, or arguments, such as “who” did “what” to “whom.” Predicate argument structure analysis deals with argument structure of verbs and adjectives. However, not only verbs and adjectives but also nouns are known to have event-hood. We thus propose a machine-learning based method for automatic argument structure analysis of Japanese event-nouns. Since there are ambiguous event-nouns in terms of event-hood, we cast the task of argument structure analysis of event-nouns into two parts: event-hood determination and argument identification. We propose to use lexico-syntactic patterns mined from large corpora for the first sub-task and to exploit argument-sharing phenomenon between predicates and event-nouns for the second sub-task.

Key Words: Predicate Argument Structure Analysis, Semantic Role Labeling, Event-nouns, Lexico-syntactic Pattern

[†]630-0192 奈良県生駒市高山町 8916-5, 奈良先端科学技術大学院大学, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan.

^{††}152-8550 東京都目黒区大岡山 2-12-1, 東京工業大学, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan.

1 はじめに

形態素解析や構文解析など自然言語処理の要素技術は成熟しつつあり、言語理解のために意味解析・談話解析といった、より高次の言語処理の研究が盛んになりつつある。特に文の意味理解のためには「誰が」「何を」「誰に」「どうした」といった要素を同定することが重要である。「誰が」「何を」「誰に」といった名詞は**項**と呼ばれ、「どうした」のような動詞を中心とした**述語**によって結びつけられる。動詞や形容詞といった述語を対象とした項構造解析は**述語項構造解析**と呼ばれ、FrameNet や PropBank といった述語項構造解析に対する資源の整備や (Gildea and Jurafsky 2002) による機械学習を用いた解析手法が登場し、近年盛んに研究されている。

述語項構造解析に関する自然言語処理の評価型ワークショップ CoNLL 2004, 2005 の開催に伴い、述語項構造解析研究はある程度の水準に達したが、深い言語理解をするためには、述語のみを対象とした事態性解析は十分でない。特に、文中の事態を指しうる表現としては、動詞や形容詞の他に名詞もあることが知られている (Grimshaw 1990)。たとえば「彼は上司の推薦で抜擢された」という文で、名詞「推薦」は「上司が彼ヲ推薦 (する)」といった事態を指す。事態とは行為や状態、出来事を指し、述語項構造と同様の項構造を考えることができる。そこで、本稿では事態を指す用法で使われていて項を持つ名詞のクラスを**事態性名詞**と呼び、事態を指す用法で使われているとき**事態性**があると定義する。本研究は、事態性名詞における項構造を抽出することを目標にしている。

事態性名詞の項構造解析とは、名詞に事態性があるとき項構造を決定し、項を同定する解析を指す。事態性とは文脈中で名詞がコト (事態¹) を指すかモノ (物体) を指すかという意味的な違いに対応する。事態性名詞の中には「レポート」のようにレポートする行為を指すのかレポートされた結果物を表すのかといった、文脈によって事態性の有無が変化する名詞がある。そこで、文脈に応じて事態性名詞に事態性があるか否か判別する処理を**事態性判別**、項構造を決定して項を同定する処理のことを**項同定**と呼ぶ。

事態性名詞の項構造解析は、述語項構造解析と同様、文中の述語の項構造を決定し、項を同定する作業の延長と位置づけることができる。英語における動詞の名詞化や日本語におけるサ変名詞など、動詞と強いつながりを持つ名詞は数多くあり、述語項構造解析の研究成果を援用して解析を行うことが期待されている。NAIST テキストコーパス (飯田, 小町, 乾, 松本 2007) によると、述語と名詞を含めた全事態中 21.1% が事態性名詞であり、述語項構造解析技術の次の発展方向として注目されている。事態性名詞の項構造解析は、情報抽出や自動要約、質問応答システム、言い換えや機械翻訳など、自然言語処理のさまざまな分野に応用できる要素技術の一つである。

¹ここで事態性というのは名詞が特定の出来事を指している場合だけではなく、総称的に使う場合も区別せず解析の対象に含める。

本研究の主な貢献は以下の2点である。

(1) **事態性判別の問題設定**：事態性判別，つまり事態を指しているかどうか曖昧性を判別する問題を設定し，事態性に関して曖昧性のない事例を用いた事態性名詞の語彙統語パターンのマイニング手法を提案した。

(2) **事態性名詞の項同定に有効な素性の提案**：事態性名詞の項構造と述語の項構造の関連性に着目し，2つの種類の素性を新たに提案した。特に動詞と格要素の共起が事態性名詞の項構造解析に有効かどうか検証し，項同定²の正解率向上に役立つことを示した。動詞と格要素の共起を用いて項同定の正解率が向上したという報告はこれまでにない。また，支援動詞構文のとき事態性名詞と述語が同じ項を共有する現象に着目し，項の対応をつけた辞書を作成して，事態性名詞の項同定に有効かどうか検証した。先行研究では明示的に支援動詞構文に関する資源を作成していないが，支援動詞辞書の整備が事態性名詞の項同定に有効であることを示した。

本論文の構成は以下のようになっている。まず2節で事態性名詞の項構造解析の先行研究について紹介する。本研究では事態性名詞の項構造解析を(1)事態性判別(2)項同定の2つの処理に分けて解く。3節でこの問題を解決するための方針について議論し，4節で事態性の曖昧性のない事例を用いた事態性名詞の語彙統語パターンのマイニング手法を提案する。5節で項同定のための動詞と格要素の共起の活用と支援動詞構文の利用について述べる。

2 関連研究

(Grimshaw 1990) は動詞と同様事態を指す名詞のことを *event nominal* (事態名詞) と呼び，*result nominal* (結果名詞)・*simple event nominal* (単純事態名詞)・*complex event nominal* (複雑事態名詞) の3つに分類した。結果名詞とは，「梅干」のように「梅を干してできたもの」という結果物を指す名詞であり，単純事態名詞とは「運動会」のように意味役割を持たない名詞である。「推薦」のように「誰が誰を誰に推薦した」という *event structure* (事態構造) を持つ複雑事態名詞³だけが項(必須格)を持つ⁴。このように，事態を指しかつ項を持つ名詞の存在は古くから知られていた。近年この現象に対する自然言語处理的観点からの関心が高まり，複数の言語でコーパスが整備されるに至った。以下で，英語・中国語・日本語における事態性名詞の項構造解析の関連研究について述べる。

² 本論文では項同定の問題のうち，項構造決定の問題は扱わない。以下，項同定は項構造が決定されたあとの項同定の問題を指す。

³ 本稿で扱う事態性名詞は Grimshaw の複雑事態名詞に相当する。

⁴ 「報告」のように結果名詞「報告書・報告物」としての語義と複雑事態名詞「出張の報告をした」としての語義両方を持つものもある。また「試験」のように文脈に応じて結果名詞・単純事態名詞・複雑事態名詞のいずれも取りうる名詞もある。

2.1 英語における事態性名詞の項構造解析

Macleod らは 1997 年から動詞の名詞化に注目し、高いカバー率の情報抽出を目的とした事態性名詞の辞書 NomLex の作成に着手した (Macleod, Meyers, Grishman, Barret, and Reeves 1997). NomLex は 2001 年に完成・公開され、NomBank プロジェクトに引き継がれる。NomBank は NomLex と同じく英語における動詞の名詞化に着目したコーパス (Meyers, Reeves, Macleod, Szekely, Zielinska, Young, and Grishman 2004c, 2004b) であり、Penn Treebank (Marcus, Santorini, and Marcinkiewicz 1993) に対し PropBank (Palmer, Kingsbury, and Gildea 2005) の仕様に従って項構造が付与されている。Meyers らは 2007 年 Penn Treebank II に対するアノテーションを終了し、NomBank 1.0 を公開した。2008 年と 2009 年の CoNLL 共通タスクでは、NomBank コーパスを用いた事態性名詞の項構造解析もタスクの一つとして行われた。

NomBank を用いた事態性名詞の項構造解析は (Jiang and Ng 2006) や (Liu and Ng 2007) がある。Jiang らは NomBank に対し、最大エントロピー法を用いた教師あり学習による項構造解析を行った。彼らは PropBank を用いた動詞に対する意味役割付与 (Semantic role labeling) において有効性が確認されている素性に加え、名詞の語幹やクラスといった事態性名詞についての意味素性や、支援動詞構文を認識するための述語との位置関係といった統語素性、そして項構造を正しく認識するための項同士の依存関係に関する大域素性を用いた。Liu らは Jiang らの用いた素性をベースに半教師あり学習手法の Alternating Structure Optimization (ASO) (Ando and Zhang 2005) を適用した。ASO は解くべき問題に関連する補助問題を作成することで経験リスク最小化を行う手法であり、彼らの研究では事態性名詞の項構造解析に有効なさまざまな補助問題が提案されている。

自然言語処理の評価型ワークショップ CoNLL 2008, 2009 では、PropBank と NomBank を用いた述語と事態性名詞に対する項構造解析の共通タスクが行われた。CoNLL 2009 には 20 チームが参加するなど、事態性名詞の項構造解析は活発に研究されている。

2.2 英語以外の言語における事態性名詞の項構造解析

中国語における事態性名詞の包括的な研究として、Xue らによる Chinese Nombank (Xue 2006a) がある。これは英語以外における初めての大規模な事態性名詞のコーパスである。

このコーパスを用いた解析として (Pradhan, Sun, Ward, Martin, and Jurafsky 2004; Xue 2006b) がある。日本語のサ変名詞と同様、中国語では動詞と動詞化された名詞は同じ表層形を持つため、動詞化された名詞は対応する動詞と共通の項構造を持つと仮定すると、動詞に関する資源を事態性名詞に流用することができる。(Xue 2006b) では、単純に動詞の事例を事態性名詞の事例に追加して実験したところ、同じ事態を指す表現であっても、動詞として使われる場合と名詞として使われる場合では語彙統語パターンが大きく違い、かえって性能が下がった、と報告している。我々は日本語を対象に、大量に自動獲得した動詞と格要素の共起を用いて事

リスク管理の必要性が強く叫ばれているが、市場の実態が把握できていないため打つ手がないのが実情。B I S が昨年春から調査の手法について検討していた。

図1 NAIST テキストコーパスの事態性名詞のアノテーション

事態性名詞の項同定を行い、大きく正解率を向上させることに成功した。また、事態性名詞に特徴的な語彙統語パターンを用いることでさらに性能を改善させた。

一方、日本語においては(黒橋 2005)によって、京都テキストコーパス第 4.0 版の一部、約 5,000 文に事態性名詞を含む名詞間の関係タグが付与された。(笹野, 河原, 黒橋 2005) は自動構築された名詞の格フレーム辞書の評価として事態性名詞の項同定を行った。彼らは事態性名詞のみについての結果を報告していないため、直接比較することはできないが、我々は事態性名詞の解析に焦点を当て、事態性判別の問題を解いた点、動詞と格要素の共起の情報を用いた点、および事態性名詞に特徴的な語彙統語パターン(支援動詞構文)を用いた点が異なる。

2.3 NAIST テキストコーパス

我々は事態性名詞の項構造解析の問題に対し、NAIST テキストコーパス(飯田他 2007)を用いた教師あり学習を行った。このコーパスでは、文章中の各事態性名詞について事態性の有無を判別し、事態性がある場合には項構造(必須格となるガ格・ヲ格・ニ格)の情報が付加されている。たとえば図1のような記事に対して、

[REL=管理(する), ガ=<外界>, ヲ=リスク]

[REL=調査(する), ガ=B I S, ヲ=実態]

のような情報が付与されている。項に関しては、「管理」のヲ格「リスク」や「調査」のガ格「B I S」のように、項が文内に出現している場合はそれが形態素単位で指示される。また、「調査」のヲ格「実態」のように、文外に出現する項でも記事内で特定できる場合はその要素が指示される。さらに「管理」のガ格のように、必須格で、かつ文内にも記事内にも出現していない場合は特別なく<外界>タグが付与されている。

現在公開されている NAIST テキストコーパス 1.4⁵ は京都テキストコーパス 3.0 全体約 4 万文に対してタグ付与されており、京都テキストコーパス 4.0 と比較して大規模な学習を行うことができるという利点がある。

(Taira, Fujita, and Nagata 2008) は NAIST テキストコーパスを対象に、決定木と SVM を用いて述語項構造解析および事態性名詞の項が存在する確率を計算して項同定を行っており、項が存在しない場合を問題として扱っているが、事態性判別のみを独立して先に解く我々の手法

⁵<http://cl.naist.jp/nldata/corpus/>

とは手順が違う。

3 事態性名詞の項構造解析へのアプローチ

事態性名詞の項構造解析を述語項構造解析と比較すると、(1) 名詞の多義性の問題 (2) 解析対象の問題 (3) 格助詞の問題がある。

1つ目の多義性の問題とは、事態性名詞の中には文脈によって事態を指す用法と指さない用法といずれも持つ名詞があるため、曖昧性を解消しなければならないという問題である。たとえば「私は公衆電話で電話をすることがめっきり減った」という文⁶では、電話はモノとしての電話であり、これ自体はなんらかの事態を指す用法ではないが、電話は「電話をする」という行為であり、「私が(誰か)ニ電話をする」という事態を指し、後者の場合のみ項が存在する。サ変名詞に限定すると、事態性のある場合は動詞としての用法を考えたときの項構造を基本的に受け継ぐ⁷ため、文脈中でのこれらの項を同定したい。

2つ目の解析単位の問題とは、項の単位の問題である。日本語の述語項構造解析では主辞が項になるため、項候補として主辞のみを解析対象に加えればよいが、事態性名詞においては主辞以外も項になりうる。たとえば、「民間支援が活性化する」では、事態性名詞支援の項のヲ格が同一文節内の主辞以外の形態素民間を指しており、主辞のみを対象にするとこのような事例を解析することができない。

(飯田他 2007)によると、述語は同一文節内に項が現れることはほとんどなく、特にヲ格とニ格においては8割以上係り受け関係にある文節の主辞が項になっている。一方、事態性名詞のヲ格とニ格はそれぞれ50.6%、43.6%が同一文節内に項を持つ。このことから、述語項構造解析で有効な統語的素性が事態性名詞の項構造解析で有効であるとは限らないことが示唆される。この問題への対策の一つとして、動詞と格要素の共起という意味情報を用いることが考えられる。

3つ目の格助詞の問題とは、述語項構造解析では明示的に格助詞(ガ・ヲ・ニ)が出現した場合はその格助詞が項同定の強い手がかりになるのに対し、事態性名詞の項構造解析においては、ほとんどの場合項は格助詞によってマークされていない、という問題である。これは述語におけるゼロ照応解析と同様の問題であり、文の構造を利用したゼロ照応解析(Iida, Inui, and Matsumoto 2006)と同じく、項同定処理に文内の構造情報を用いることが有効であると考えられる。特に支援動詞構文のとき、事態性名詞と述語は項を共有するので、明示的な格助詞を伴う述語の項構造を解析に用いることができる。

さて、事態性名詞の項構造解析をするに当たり、我々はまず(1)の問題に対応する事態性判

⁶以下事態性のない用例は下線、事態性のある用例は四角で囲んで示す。

⁷「動き」「薦め」などの和語動詞由来の事態性名詞も同様に扱うことができる。また、「運動会」のように事態を指すが動詞と直接の結びつきがない事態性名詞は今回扱わない。

別を行い、その後事態性のある事態性名詞に限って項構造解析を行うという手順で問題を解く。

このようにモデルを分ける利点は、事態性判別は語義曖昧性解消の問題であり、項同定と別の素性を用いた解析が有効だと考えられ、モデルを分けることにより事態性判別に特徴的な統語素性を用いることができる点である⁸。

事態性判別は文中に出現する事態性名詞を事態性あり／なしの2クラスに分類する問題なので、事態性に関する曖昧性のない事例を用いて教師なしのパターンマイニングを行うことができる。語義曖昧性解消タスクにおいては曖昧性のある単語周辺の単語の分布が有効な素性であることが知られており、大規模なコーパスから文脈素性を学習することが本タスクにおいても有効であることが期待される。そこで、本論文ではこの方法に従って事態性判別のタスクと項同定のタスクを分けて扱う手法を提案する。

4 事態性判別

事態性名詞には、「リスク管理」や「彼の決断」のように、事態性名詞と同一文節内や係り受け関係にある文節内に項が存在することが多く、こういった名詞の出現パターンを利用することによって事態性判別の精度が向上すると考えられる。

そこで、名詞の出現パターンを捉えるための手段として BACT (Kudo and Matsumoto 2004)⁹ を用いて事態性のある名詞と事態性のない名詞との出現パターンを学習することを考えた。BACT は木構造の訓練事例を学習させることによってブースティングで訓練事例の判別に効果の高い重み付き部分木を順次選択するアルゴリズムであり、文の構造を木構造に変換して素性として入れることによって、訓練事例の判別に効果が高い構造をルールとして学習できる。

4.1 事態性名詞の語彙統語パターンの学習

事態性名詞の語彙統語パターンの獲得には、名詞の出現する前後3形態素、名詞の出現する文節および係り元の文節の形態素列を木構造にして分類器を作成した。事態性に関する曖昧性がない事例として、日本語語彙大系(池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林 1997) にサ変動詞として登録されている用言のうち、一般名詞意味属性体系の「名詞-抽象-事-{人間活動, 事象}」ノードの下にあり、かつそれ以外のノードの下にない2,253名詞を曖昧性のない事態性名詞として正例にした。ランダムに200個サンプリングして調べたところ、97%がサ変名詞、残りの3%が動詞由来の事態性名詞であった。また、一般名詞として登録されている名詞のうち「名詞-具体」ノードの下にあり、かつそれ以外のノードの下にない名詞と固有名詞合わせ

⁸述語項構造解析においては項同定と項識別とで有効な素性が大きく異なり、場合によっては項同定に有効な素性が項識別ではノイズになりうる、という指摘(Pradhan, Hacioglu, Krugler, Ward, Martin, and Jurafsky 2005)もある。

⁹<http://chasen.org/~taku/software/bact/>

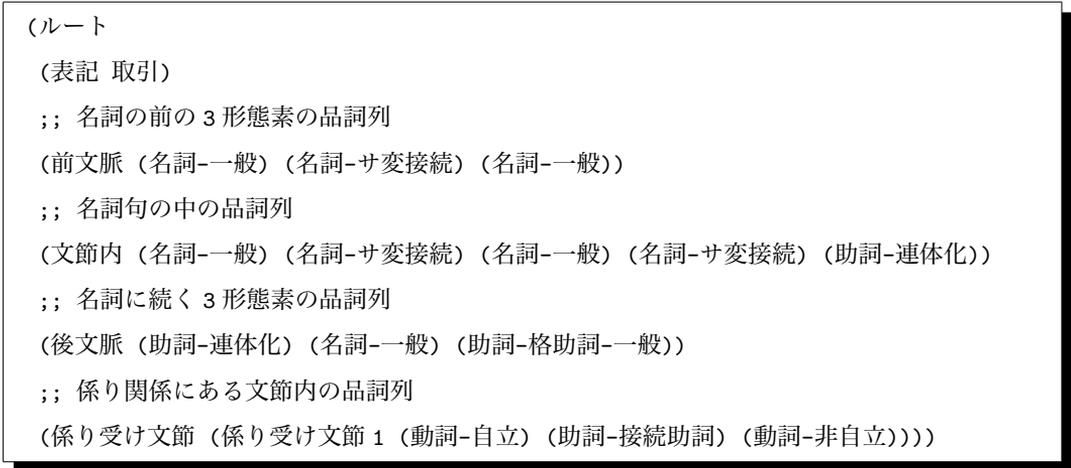


図 2 「商品取引」のパターンの例

表 1 事態性判別に有効な素性として獲得したパターンの例

事態性ありの判定に効果が高いパターン	スコア	事態性なしの判定に効果が高いパターン	スコア
同一文節中にサ変名詞+サ変名詞がある	1.13	後ろにサ変名詞がある	-0.92
後ろに助詞+サ変名詞が続く	1.01	前にサ変名詞がある	-0.54
後ろに名詞の接尾辞がある	0.61	前にサ変名詞, 後ろに助詞がある	-0.50

て 194,098 名詞を曖昧性のない非事態性名詞として負例にした。ランダムに 200 個サンプリングして調べたところ、16.5%が一般名詞、73.5%が固有名詞で非事態性名詞であった。たとえば「商品取引」の出現パターンは図 2 のような木構造になり、正例として訓練事例に追加する。

学習には新聞記事約 1 ヶ月分 (毎日新聞社 2002) (正例:117,581 事例, 負例:282,419 事例) を使用し、正例および負例を分類するに当たっての重みが高いルール 6 つを獲得し、表 1 に示した。重みの絶対値が大きければ大きいほど、事態性判別に効果が高いと考えられる。

4.2 実験

事態性の判別には語義曖昧性解消で最もよい性能を示している Support Vector Machines (Vapnik 1998) (実装は TinySVM¹⁰) を用い、文脈に応じた事態性名詞の事態性を学習した。多項式 2 次カーネルを使用し、その他のパラメータはデフォルト値を用いた。評価は精度・再現率・F 値 (精度と再現率の調和平均) で行い、10 分割交差検定によって事態性の有無の判別性能を

¹⁰<http://www.chasen.org/~taku/software/TinySVM>

表 2 事態性判別に用いた素性

文節内素性
事態性名詞の分類語彙表 (国立国語研究所 2004) 中での分類項目
事態性名詞が複合名詞であったとき, 各名詞についての分類語彙表中の分類項目の上位 4 桁
文節外素性
事態性名詞の前後 1 文節の形態素列
「A の B」(例: 税の軽減) の形のとき, A に当たる名詞の分類語彙表中の分類項目の上位 4 桁
名詞の語彙統語パターン
同一文節中にサ変名詞+サ変名詞がある
後ろにサ変名詞がある
後ろに助詞+サ変名詞が続く
前にサ変名詞がある
後ろに名詞の接尾辞がある
前にサ変名詞, 後ろに助詞がある

表 3 事態性判別実験結果

	精度	再現率	F 値
最頻出語義 (ベースライン)	60.4%	88.2%	71.7%
語彙統語パターンなし	72.2%	58.7%	64.8%
語彙統語パターンあり	73.3%	80.2%	76.6%

見た, 評価事例には NAIST テキストコーパスから新聞記事 80 記事 (800 文) を用いた, 含まれる事態性名詞は 1,237 個 (うち 590 個が事態性ありの事例) あった. ベースラインには各名詞に対してコーパス中で最も頻度が高い語義を正解としたモデルを用い, BACT によってマイニングした名詞の出現パターンを用いたモデルと比較した. 使用した素性は表 2 にまとめた.

教師なしでマイニングした語彙統語パターンの有効性を示すため, 表 3 に事態性判別の実験結果を示した. 提案手法はベースラインより再現率は劣るものの, 精度を大幅に向上させ, F 値で約 5%改善することができた. 名詞の語彙統語パターンを用いない場合 (表 2 の文節内素性と文節外素性のみを用いた場合), ベースラインより精度は上がるものの再現率は大幅に低下し, 文節内の品詞情報など局所的な文法知識が事態性名詞の判別に効果が高いことを示している.

4.3 エラー分析

事態性がある事例にも関わらず事態性がないと判別を誤った事例に次のようなものがあった。

- 「野良黒山の会」のリーダー、木場将弘さん方では、妻の和枝さんらが現地と電話のやり取りを続けた。
- 自民党の渡辺美智雄元副総理・外相は四日、宇都宮市で講演、七月の参院選について「社会党と連合が独占している二十三選挙区でも自民党から候補者を擁立すべきだ。連立政権は政権。選挙は選挙だ」と述べ、選挙協力よりも独自候補擁立を優先すべきだとの考えを示した。

これらの事例はいずれも文内に項があるにも関わらず事態性判別を誤った事例であり、事態性判別に項同定の情報を用いていないために判別に失敗している。これらの事例に正解するためには、事態性判別と項同定を同時に最適化する必要がある。解決策として、(Iida et al. 2006) の提案する述語のゼロ照応解析で用いられている探索先行分類モデルを適用し、まず最尤の項候補を求め項同定を先に行った上で事態性判別の一素性として用いる方法や、Markov Logic Networkを用いて事態性判別と項同定を同時に行う (Meza-Ruiz and Riedel 2009) といった方法が考えられるが、どのようにするべきかは今後の課題である。

5 事態性名詞の項同定に有効な素性の検討

3節で述べたように、事態性名詞の項同定は述語の項同定と異なり、述語項構造解析で有効であった文法的素性が必ずしも有効であるとは限らない。そこで、動詞と格要素の共起といった意味的素性を用いることを提案する。また、事態性名詞の項は述語の項と違い格助詞を伴わないという問題に対し、支援動詞構文において事態性名詞と述語が項を共有する現象に着目し、支援動詞辞書を作ることにより、項の対応関係を認識する手法について述べる。

5.1 動詞と格要素の共起の利用

事態性名詞のうち、対応する動詞がある名詞、たとえばサ変名詞 (例：推薦) や動詞由来の名詞 (例：動き) の項構造は動詞の項構造と類似しており、事態性名詞の項同定に動詞に関する知識が有効であろうと考えられる。

述語においてはヲ格の 84%とニ格の 88%が係り受けの関係にある文節に項を持つが、事態性名詞においてはそれぞれヲ格の 31%、ニ格の 22%しか係り受け関係の文節に項が存在しない (飯田他 2007)。それに加え、述語では項は普通格助詞ヲやニを伴って出現するため、表層の格助詞が大きな手がかりとなるが、事態性名詞の項は明示的に格助詞を伴わない。

この問題に対処するため、まずサ変名詞とそれに対応する動詞が意味的には共通の項を持つ

ことを仮定し、動詞と格要素の共起情報を事態性名詞の項同定に用いる¹¹。この手法の利点は、係り受け解析器を用いることで大規模なラベルなしデータから自動的に獲得した共起を獲得し、完全に教師なしに行うことができる点である。

動詞と格要素の共起のモデル化は(藤田, 乾, 松本 2004)のモデルに従った。このモデルは、名詞 n が格助詞 c を介して動詞 v に係っているときの共起確率 $P(\langle v, c, n \rangle)$ を推定するため、 $\langle v, c, n \rangle$ を $\langle v, c \rangle$ と n の共起と見なす。しかしながら、一般に動詞と格要素の共起事例は非常に疎なので、pLSI(Hoffman 1999)を用いてスムージングを行う。

$$P(\langle v, c, n \rangle) = \sum_{z \in Z} P(\langle v, c \rangle | z) P(n | z) P(z)$$

Z は共起に関する潜在的な意味クラスを指す確率変数で、確率分布を用いて単語行列を $|Z|$ 次元に圧縮していることに相当する。 $P(\langle v, c \rangle | z), P(n | z), P(z)$ は EM アルゴリズムで求められる。

共起尺度としては自己相互情報量 (Hindle 1990) を用いる。

$$PMI(\langle v, c \rangle, n) = \log_2 \frac{P(\langle v, c, n \rangle)}{P(\langle v, c \rangle)P(n)}$$

共起尺度の利用に関しては、相対的な共起の強さを用いて項らしさを判定する。そこで、これらの共起尺度により計算されたスコアは、最尤候補に対する相対値として素性に取り込む。具体的には、項候補となる名詞句をペアに対し、それらの共起スコアの差（実数値、およびとそれを離散化したもの）を素性として用いる。

本論文では藤田らに倣い、新聞記事延べ 19 年分のコーパスから係り受け解析結果を抽出し、 $|Z| = 1000$ としてモデルを作成した。

事態性名詞と項の共起の分布を調べた予備実験において、固有名詞を中心に、動詞と格要素の共起モデル中出现しなかった事例が全体の 18% あった。そこで、共起スコアを計算することができなかった名詞と、共起スコアが負であった名詞に関して、CaboCha の出力する固有表現ラベルおよび ChaSen で品詞が固有名詞であると判定された名詞に対しては、固有表現クラスで推定した共起スコアを用いてスムージングした¹²。

5.2 支援動詞辞書の作成

支援動詞構文は述語と事態性名詞が項を共有する統語パターンであり、述語に係る名詞句の格助詞の情報を使うことによって性能を向上させることができると考えられる。そこで、我々は事態性名詞と述語の間で項の対応がついた辞書（支援動詞辞書）を作成し、支援動詞構文の認識に用いた。事態性名詞の 21.7% は支援動詞構文で使われており、述語項構造解析で有効な

¹¹ 動詞由来の名詞にもこの手法は拡張可能である。

¹² 固有表現クラスによるスムージングは後述する項同定の予備実験で常に高い精度だったため、以降の実験ではスムージングを用いた結果のみ報告した。

格助詞に関する情報を活用できると予想される。

NomBank をコーパスに用いた分析として、(Meyers, Reeves, and Macleod 2004a) は、名詞化された動詞の多くは支援動詞構文にあり、主動詞と項を共有することを指摘した¹³。(Jiang and Ng 2006) も同じ現象に着目し、機械学習の素性として支援動詞構文を用いることを提案した。

しかしながら、支援動詞構文となる述語は限られており、人手で書き尽くすことができるため、我々は事態性名詞と述語の間で項を共有するようなペアについて辞書を作成し、どのような項の対応関係になっているか、という情報を付与して用いることにした。

たとえば、文「太郎が花子に電話をした」において、述語項構造は [REL=する (した), ガ=太郎, ヲ=電話, ニ=花子] であり、事態性名詞の項構造は [REL=電話 (する), ガ=太郎] である。この場合、述語「する」と事態性名詞「電話」は項「太郎」をそれぞれガ格で共有する。表5はこの場合支援動詞辞書のどのエントリにマッチしているのか示した。また、「彼が彼女に勉強を教える」という文では、述語「教える」は事態性名詞と項を共有しているが、述語の二格「彼女」は事態性名詞ではガ格となり（「彼女」が「勉強 (する)」）、格の交替が起きる。表6には格交替が起きる場合の項の対応を示した。

NAIST テキストコーパスには異なり 2,173 個、延べ 8,190 個の支援動詞構文の事例が見られた。頻度上位 10 件の支援動詞構文を表 4 にまとめた。ここから分かるように、実際のコーパスに出現する支援動詞は、いわゆる機能動詞結合で用いられる機能動詞より広い概念である。たとえば、「太郎が電話を続ける」においては、「太郎が電話する」という事態が継続していることを示し、「続ける」は継続の意味を失っていない。このように、支援動詞構文にある事態性名詞と述語の示す事態の関係は、テンスやモダリティの付加などがあるが、このような動詞がいくつあり、どのような関係がありうるかは今後の研究課題である。

支援動詞辞書を作成するに当たって、Web 5 億文コーパス (Kawahara and Kurohashi 2006) から事態性名詞が直接係っている述語とその項を 200 万事例抽出した¹⁴。そのうち、頻度上位 2,000 件 (抽出された 200 万事例中、延べ 80% をカバー) を対象に、事態性名詞が述語と項を共有しているかどうかを判断した。具体的には、述語と事態性名詞のそれぞれの格 (ガ・ヲ・ニ) について、事態性名詞の含まれる格と、共有している場合には共有している格の対応関係の情報を付与した。項対応の判断に関しては、Web コーパス中から取得した周辺文脈を補助的に用い、周辺文脈頻度上位 5 件中 3 件以上で成り立つ場合のみ対応関係を付与した。

5.3 実験

事態性名詞に特徴的な語彙統語パターンの情報を用いることで、項同定の性能が向上することを示す実験を行った。述語項構造解析のトーナメントモデル (Iida, Inui, and Matsumoto 2005)

¹³ここでは彼らの定義に従い、支援動詞とは最低 2 つの項 NP₁ と XP₂ を取り、XP₂ が NP₁ の主辞の項となっている場合を指す。

¹⁴依存構造解析は KNP (<http://nlp.kuee.kyoto-u.ac.jp/nl-resource/knp.html>) による。

表 4 支援動詞構文の頻度上位 10 件. S, N, E はそれぞれ共有された項, 共有されていない名詞句, 事態性名詞を表す.

支援動詞構文	例	カバー率
S ガ E ヲ する	太郎が電話をする	14.4%
S ガ E ヲ 行う	太郎が演説を行う	3.4%
S ガ E ニ する	太郎が参考にする	1.7%
S ガ E ヲ N ニ する	太郎が約束を反故にする	1.4%
S ガ E ヲ 受ける	太郎が注目を受ける	1.4%
S ガ E ヲ 続ける	太郎が電話を続ける	1.3%
S ガ E ヲ 進める	太郎が研究を進める	1.3%
S ガ E ヲ 目指す	太郎が出馬を目指す	1.2%
N ガ E ヲ S ニ 求める	次郎が推薦を太郎に求める	1.0%
S ガ E ヲ 図る	太郎が自殺を図る	0.9%

表 5 「する」の支援動詞辞書項目と「太郎が花子に電話をする」の項対応. E には共通の事態性名詞が, S には共有される項が入る. このエントリは二格の共有に関する情報は入っていない.

	REL	ガ	ヲ	ニ
述語	する	S (太郎)	E (電話)	
事態性名詞	E (電話)	S (太郎)		

表 6 「教える」の支援動詞辞書項目と「彼が彼女に勉強を教える」の項対応. E には共通の事態性名詞が, S には共有される項が入る. 二格とガ格で格の交替が起きている.

	REL	ガ	ヲ	ニ
述語	教える		E (勉強)	S (彼女)
事態性名詞	E (勉強)	S (彼女)		

を事態性名詞の項構造解析に適応し, 事態性名詞の項同定システムを作成した. トーナメントモデルに用いる三つ組は事態性名詞, 正解の項, そして項候補である. 各三つ組に対し, システムは訓練時どちらの候補が勝つか, そしてそれを特徴づける素性を学習する. テスト時には項の候補を順に試し, トーナメントで最後まで勝ち残った候補を予測する項として提示する.

今回ベースラインシステムと提案する 3 つのモデルを比較した.

ベースラインシステムで用いた素性は表 7 に示した. 使用した素性は (Iida et al. 2006) に準

表 7 項同定に用いた素性リスト

語彙素性	位置素性
表層形	事態性名詞と項の距離 (形態素単位)
文法素性	文節内の先頭からの位置 (形態素単位)
品詞	文頭にあるかどうか (0/1)
* 同一文節内の他の形態素の品詞	文末にあるかどうか (0/1)
名詞が定名詞かどうか (0/1)	* 文節頭にあるかどうか (0/1)
名詞が指示代名詞かどうか (0/1)	* 文節末にあるかどうか (0/1)
助詞 (表層形)	事態性名詞が項より前にあるか (0/1)
意味素性	項が事態性名詞より前にあるか (0/1)
新聞記事から計算した共起スコアの値 (実数)	事態性名詞が直接項に係っているか (0/1)
Web 5 億文から計算した共起スコアの値 (実数)	* 事態性名詞から項までのパスがあるか (0/1)
* 名詞句のペアの新聞共起スコアの差 (実数)	項が直接事態性名詞に係っているか (0/1)
* 名詞句のペアの Web 共起スコアの差 (実数)	* 項から事態性名詞までのパスがあるか (0/1)
日本語語彙大系の構文パターン対による選択制限	* 事態性名詞の係り先の述語が機能動詞か (0/1)
代名詞のタイプ	* 項が機能動詞に係っているか (0/1)
固有名詞のタイプ	* 述語と事態性名詞が項を共有しているか (0/1)
	引用の中か (0/1)

じ、事態性名詞で計算することのできない素性¹⁵ は使用しなかった。

事態性名詞の項同定は形態素単位で行うため、文節内の位置の素性、同一文節内の他の形態素の品詞列の素性を追加した。また、(村木 1990) に掲載されている機能動詞 128 表現を用い、事態性名詞が機能動詞に係っているかどうかの素性と、機能動詞に係っている場合に機能動詞と項が係り関係にあるかどうかの素性を追加した。本論文で新しく導入された素性にはアスタリスクを付した。

また、実験に用いたデータは NAIST テキストコーパスから新聞記事 1 日分 (137 記事, 延べ 847 個の事態性名詞, 97 個の支援動詞構文) を訓練事例, それとは異なる 1 日分 (150 記事, 延べ 722 個の事態性名詞, 113 個の支援動詞構文) を評価事例に用いた。実験に用いた機械学習器は Support Vector Machines (Vapnik 1998) であり, 実装は TinySVM¹⁶, 多項式 2 次カーネルでパラメータはデフォルト値を使用した。

支援動詞辞書と共起モデルの有効性を示すため, 各モデルについてそれぞれの格における正解率を求め, 表 8 に示した。ガ・ヲ・ニ格を対象に, それぞれ文内に項を持つ事例のみを訓練・テストに用い, 項同定の正解率で評価した。「+ 共起モデル」はベースラインモデルに加え, 共起のスコアを実数値の素性として Support Vector Machines に適用したものであり, 「+ 支援動

¹⁵ 事態性名詞にはヴォイスと助動詞は定めることができない。また, センタリング理論に基づいた素性は用いていない。EDR を用いた素性は予備実験により正解率が低下したので削除した。

¹⁶ <http://chasen.org/~taku/Software/TinySVM/>

表 8 事態性名詞の項同定タスクの正解率

素性	ガ	ヲ	ニ
ベースライン	60.5	79.7	73.0
+ 支援動詞辞書	64.2	78.0	71.4
+ 共起モデル	67.1	80.1	74.6
+ 支援動詞辞書 + 共起モデル	68.3	80.1	74.6

詞辞書」はテスト事例で支援動詞辞書にマッチする項目がある場合、SVM によるトーナメントモデルを用いず、辞書に対応する述語の項をそのままシステムの出力とした¹⁷ものである。

共起モデルを用いた手法はどの格においてもベースラインより高い正解率を示した。また、支援動詞辞書を共起モデルと組み合わせたモデルはヲ格とニ格においてベースラインよりわずかながら高い正解率であった。

5.4 議論

支援動詞辞書を用いてもっとも効果が高かったのはガ格であるが、これは項の共有があった格のうち 92% はガ格であるためだと考えられる。テスト事例で支援動詞構文にあったもののうち、項と述語が係り受け関係にある事例は 38 事例であり、75 事例 (66%) は省略解析の情報をタグ付きコーパスから取得できたことが精度向上に貢献していると考えられる。ここで示した精度は述語項構造が正しく解析できた場合の上限値であり、述語項構造を自動解析した場合は精度が低下することが考えられるが、実際に省略解析をどのように行い、どのような結果を得るかは解析に採用する解析モデルに依存する。

一方、ヲ格とニ格において、支援動詞辞書を用いるとベースラインより性能が悪くなった。ヲ格で支援動詞辞書の効果がなかった理由としては、ヲ格の 90% が既に事態性名詞と同一文節内にあるか係り受け関係にある文節に存在し、明示的に項の交替関係をモデル化する必要がなかったことが推測される。また、支援動詞の項の共有情報を素性として用いた実験を行ったところ、ガ格とニ格ではベースラインと正解率は変わらず、ヲ格では正解率の低下が見られた。

支援動詞辞書の効果が格によって異なっている原因を調べるため、支援動詞辞書のみシステムの性能を求めた。Web コーパスから作成した支援動詞辞書のカバー率は、新聞記事分野で作成した事態タグつきコーパスに対して 49% であった。支援動詞辞書のパターン対にマッチする事態性名詞に対し、支援動詞辞書のみを用いた項同定システムの性能を計ると精度は 0.72、再現率は 0.35 であった。ヲ格とニ格に対してはこの精度はベースラインの正解率を下回ってい

¹⁷述語の項はコーパスに人手で付与された正解データを用いる。

る。そのため、支援動詞辞書の情報を用いるとヲ格とニ格に関しては、正解率が低下する可能性があることが分かった。実際、表8では、まさにヲ格とニ格では支援動詞辞書を使うことによって正解率が低下している。この問題への対処法としては、ヲ格とニ格に対しては、精度の高い支援動詞構文のみ用いることが考えられる。どのようにして効率的に支援動詞辞書を構築するかは今後の課題である。

提案手法の典型的な誤り事例は、局所的な項を適切に同定できないという誤りである。たとえば「太郎が次郎の連勝を止めた」という例文では正しい項構造は [REL=連勝, ガ=次郎] だが、システムは [REL=連勝, ガ=太郎] を出力した。これはこの事例が「X を止める」という辞書項目にマッチし、同一名詞句内にある候補「次郎」を解析しなかったためである。この問題に対処するには、局所的な候補から順番に項を探し、ふさわしい候補が見つからなかった場合に探索範囲を広げていく、といった階層的なモデルを用いることが考えられる。

6 おわりに

本論文で、名詞句の語彙統語パターンを用いた事態性名詞の項構造解析について述べた。本論文では項構造解析のタスクを2つに分け、マイニングした語彙統語パターンを用いた事態性名詞の事態性判別手法を提案した。また、動詞と格要素の共起と事態性名詞に特徴的な語彙統語パターンが事態性名詞の項同定に有効であることを示した。

提案手法では事態性判別は精度76.6%、再現率79.6%で行うことができる。また、項同定も文内の項ではガ格・ヲ格・ニ格それぞれ68.3%・80.1%・74.6%の正解率で解析でき、ある程度実用的に項構造解析を行うことができた。まだ高速化・精度改善の余地はあるが、本研究をベースに情報抽出などの応用に用いる下地ができたと考えている。

今後は形態素解析から固有表現抽出、係り受け解析までを含めて同時に最適化を行なう手法の研究や、項の間の依存関係を考慮した項同定モデルの研究が考えられる。

謝 辞

本研究の一部は科研費特定領域研究「代表性を有する大規模日本語書き言葉コーパスの構築」の助成を受けたものである。Web から取得した5億文データを使用させてくださった河原大輔氏にと、新聞記事から抽出した動詞と格要素の共起モデルおよび自己相互情報量の計算プログラムを利用させてくださった藤田篤氏にお礼申し上げる。また、日本語述語項構造解析に関して有益なコメントをくださった平博順氏に感謝する。

参考文献

- Ando, R. K. and Zhang, T. (2005). “A Framework for Learning Predictive Structures from Multiple Tasks and Unlabeled Data.” *Journal of Machine Learning Research*, **6**, pp. 1817–1853.
- Gildea, D. and Jurafsky, D. (2002). “Automatic Labeling of Semantic Roles.” *Computational Linguistics*, **28** (3), pp. 245–288.
- Grimshaw, J. (1990). *Argument Structure*. MIT Press.
- Hindle, D. (1990). “Noun Classification from Predicate Argument Astructures.” In *Proceedings of the 28th Annual Meeting of the ACL*, pp. 268–275.
- Hoffman, T. (1999). “Probabilistic latent semantic indexing.” In *Proceedings of the 22nd Annual ACM Conference on Research and Development in Information Retrieval*, pp. 50–57.
- Iida, R., Inui, K., and Matsumoto, Y. (2005). “Anaphora resolution by antecedent identification followed by anaphoricity determination.” *ACM Transactions on Asian Language Information Processing (TALIP)*, **4** (4), pp. 417–434.
- Iida, R., Inui, K., and Matsumoto, Y. (2006). “Exploiting Syntactic Patterns as Clues in Zero-Anaphora Resolution.” In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL (COLING-ACL)*, pp. 625–632.
- Jiang, Z. P. and Ng, H. T. (2006). “Semantic Role Labeling of NomBank: A Maximum Entropy Approach.” In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Kawahara, D. and Kurohashi, S. (2006). “Case Frame Compilation from the Web using High-Performance Computing.” In *Proceedings of the fifth International Conference on Language Resource and Evaluation (LREC)*, pp. 1344–1347.
- Kudo, T. and Matsumoto, Y. (2004). “Boosting Algorithm for Classification of Semi-Structured Text.” In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 301–308.
- Liu, C. and Ng, H. T. (2007). “Learning Predictive Structures for Semantic Role Labeling of NomBank.” In *Proceedings of the 46th Annual Meeting of the ACL*.
- Macleod, C., Meyers, A., Grishman, R., Barret, L., and Reeves, R. (1997). “Designing a Dictionary of Derived Nominals.” In *Proceedings of Recent Advances in Natural Language Processing*.
- Marcus, M. P., Santorini, B., and Marcinkiewicz, M. A. (1993). “Building a Large Annotated Corpus of English: The Penn Treebank.” *Computational Linguistics*, **19** (2), pp. 313–330.
- Meyers, A., Reeves, R., and Macleod, C. (2004a). “NP-External Arguments: A Study of Argu-

- ment Sharing in English.” In *Proceedings of the ACL 2004 Workshop on Multiword Expressions: Integrating Processing*, pp. 96–103.
- Meyers, A., Reeves, R., Macleod, C., Szekely, R., Zielinska, V., Young, B., and Grishman, R. (2004b). “Annotating Noun Argument Structure for NomBank.” In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC)*, pp. 803–806.
- Meyers, A., Reeves, R., Macleod, C., Szekely, R., Zielinska, V., Young, B., and Grishman, R. (2004c). “The NomBank Project: An Interim Report.” In *Proceedings of the HLT/NAACL 2004 Workshop Frontiers in Corpus Annotation*, pp. 24–31.
- Meza-Ruiz, I. and Riedel, S. (2009). “Jointly Identifying Predicates, Arguments and Senses using Markov Logic.” In *Proceedings of Humal Language Technologies: The 2009 Annual Conference on the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2009)*, pp. i155–163.
- Palmer, M., Kingsbury, P., and Gildea, D. (2005). “The Proposition Bank: An Annotated Corpus of Semantic Roles.” *Computational Linguistics*, **31** (1), pp. 71–106.
- Pradhan, S., Hacioglu, K., Krugler, V., Ward, W., Martin, J. H., and Jurafsky, D. (2005). “Support Vector Learning for Semantic Argument Classification.” *Machine Learning*, **60** (1-3), pp. 11–39.
- Pradhan, S., Sun, H., Ward, W., Martin, J. H., and Jurafsky, D. (2004). “Parsing Arguments of Nominalizations in English and Chinese.” In *Proceedings of the HLT/NAACL*.
- Taira, H., Fujita, S., and Nagata, M. (2008). “A Japanese Predicate Argument Structure Analysis using Decision Lists.” In *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP 2008)*, pp. 522–531.
- Vapnik, V. N. (1998). *The Statistical Learning Theory*. Springer.
- Xue, N. (2006a). “Annotating the Predicate-Argument Structure of Chinese Nominalizations.” In *Proceedings of the fifth International Conference on Language Resource and Evaluation (LREC)*, pp. 1382–1387.
- Xue, N. (2006b). “Semantic Role Labeling of Nominalized Predicates in Chinese.” In *Proceedings of the HLT-NAACL*, pp. 431–438.
- 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳文, 林良彦 (編) (1997). 日本語語彙大系. 岩波書店.
- 国立国語研究所 (2004). 分類語彙表. 大日本図書株式会社.
- 村木新二郎 (1990). 日本語動詞の諸相. ひつじ書房.
- 黒橋禎夫 (2005). 京都テキストコーパス Version 4.0. <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/corpus.html>.

- 藤田篤, 乾健太郎, 松本裕治 (2004). “自動生成された言い換え文における不適格な動詞格構造の検出.” 情報処理学会論文誌, **45** (4), pp. 1176–1187.
- 毎日新聞社 (2002). 毎日新聞. 毎日新聞社.
- 飯田龍, 小町守, 乾健太郎, 松本裕治 (2007). “NAIST テキストコーパス: 述語項構造と共参照関係のアノテーション.” 情報処理学会研究会報告 (自然言語処理研究会), pp. 71–78. NL-177-10.
- 笹野遼平, 河原大輔, 黒橋禎夫 (2005). “名詞格フレーム辞書の自動構築とそれを用いた名詞句の関係解析.” 自然言語処理, **12** (3), pp. 129–144.

略歴

- 小町 守**：2005 年東京大学教養学部基礎科学科科学史・科学哲学分科卒。2007 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。博士後期課程に進学。修士（工学）。2008 年より日本学術振興会特別研究員（DC2）。大規模なコーパスを用いた意味解析に関心がある。言語処理学会第 14 回年次大会最優秀発表賞受賞。人工知能学会，情報処理学会，ACL 各会員。
- 飯田 龍**：2007 年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。同年より奈良先端科学技術大学院大学情報科学研究科特任助教。2008 年 12 月より東京工業大学大学院情報理工学研究科助教。現在に至る。博士（工学）。自然言語処理の研究に従事。情報処理学会員。
- 乾 健太郎**：1995 年東京工業大学大学院情報理工学研究科博士課程修了。博士（工学）。同研究科助手。九州工業大学情報工学部助教授を経て、2002 年より奈良先端科学技術大学院大学情報科学研究科助教授。現在同研究科准教授。情報通信研究機構有期研究員を兼任。自然言語処理の研究に従事。Computational Linguistics 編集委員，情報処理学会，人工知能学会，ACL 各会員。
- 松本 裕治**：1977 年京都大学工学部情報工学科卒。1979 年同大学大学院工学研究科修士課程情報工学専攻修了。同年電子技術総合研究所入所。1984～85 年英国インペリアルカレッジ客員研究員。1985～87 年（財）新世代コンピュータ技術開発機構に出向。京都大学助教授を経て、1993 年より奈良先端科学技術大学院大学教授。現在に至る。工学博士。専門は自然言語処理。人工知能学会，日本ソフトウェア科学会，情報処理学会，認知科学会，AAAI, ACL, ACM 各会員。

(2009 年 6 月 8 日 受付)

(2009 年 10 月 8 日 再受付)

(2009 年 10 月 28 日 採録)