

令和 6 年 6 月 10 日現在

機関番号：12605

研究種目：国際共同研究加速基金（国際共同研究強化）

研究期間：2018～2023

課題番号：17KK0002

研究課題名（和文）語義曖昧性解消技術と領域適応手法を利用した情報抽出手法の研究開発

研究課題名（英文）Research and development of information extraction methods using word sense disambiguation and domain adaptation

研究代表者

古宮 嘉那子（Komiya, Kanako）

東京農工大学・工学（系）研究科（研究院）・准教授

研究者番号：10592339

交付決定額（研究期間全体）：（直接経費） 7,800,000円

渡航期間：12ヶ月

研究成果の概要（和文）：日本語の科学技術論文から「問題」（problem: 困ったことという意味の方で、task: お題ではない）の内容を抽出する研究を行った。英語の科学技術論文から同様のことを行った論文を参考に始めたが、日本語において問題を表す表現が複雑であることから、その表現形式についてアノテーションルールを定め、言語学的な定義を行った。また、このルールに沿って、文中の「問題」が指す問題内容がその文に含まれているか否かについてのアノテーションを行い、様々な手法で分類実験を行った。

研究成果の学術的意義や社会的意義

英語と比較して、日本語の論文における問題内容の書かれ方について分析を行った。英語論文ではThe problem is X.の書かれ方で書かれている問題内容のみを扱っていたが、日本語では、コピュラ的な表現「Xが問題だ」以外にも修飾的な表現「Xという問題」のような表現が多くみられることが分かった。これらを踏まえて、問題内容のアノテーションルールを策定し、コーパスを作成した。この際に、問題内容は入れ子構造になっていることがあること、問題内容を示すのは、文のことも単語やフレーズのこともあること、指し示す問題内容の粒度にばらつきがあることなどを分析し、ルールに反映した。

研究成果の概要（英文）：We extracted the statements of "problems" (meaning something problematic, not tasks) from Japanese scientific and technical papers. We started by referring to a paper that did the same thing from English scientific and technical papers. However, because the expressions used to describe problems are complex in Japanese, we established annotation rules for the forms of expression and defined them linguistically. Following these rules, we also annotated whether or not the problem statements referred to by 'problem' in a sentence were included in that sentence, and conducted classification experiments using various methods.

研究分野：自然言語処理

キーワード：問題抽出 アノテーション 科学技術論文 語義曖昧性解消 BERT

1. 研究開始当初の背景

申請者は基課題において、日本語の語義曖昧性解消に関する研究を行ってきた。これは、文中の単語の意味をコンピュータに理解させるという自然言語処理の基礎研究の一環である。また、語義曖昧性解消タスクを対象タスクとした、ドメイン適応の研究を長らくおこなってきた。ドメイン適応とは、学習データとは異なった性質を持つデータを処理対象とする際に、性能を上げるための技術である。

一方、言語処理によって自動的に得られた文中の単語の意味が、情報工学分野の論文から論文中で解決されるべきとされる問題点とその解決法を抽出するという情報抽出のタスクにおいて、手がかりとして有効であったという先行研究があり (Heffernan et al., 2017)、これが共同研究先の Cambridge 大学の Teufel 教授の研究室における研究であった。

(Heffernan et al., 2017) は、「The problem is X.」という形の文を論文から抽出して、解決すべき問題点であるかを判定している。申請当時 (2017 年) はコロナ禍前であり、また、大規模事前学習モデルが登場する前であったため、(Heffernan et al., 2017) では統計的機械学習モデルの種類である、Support Vector Machine を利用して解決されるべきとされる問題点が書かれている文であるか否かを判定する研究を行っていた。また、対象としている論文は、英語の自然言語処理の論文に限られていた。この論文において、Heffernan らは problem であるとされている文を抽出しているが、problem には(1)困ったこと (抽出対象となる問題) と(2)クイズなどのお題 (抽出対象ではない問題) という二つの意味があり、この違いを加味できれば、解決すべき課題か否かの分類性能が向上するのではないかと期待があった。

2. 研究の目的

申請者は (Heffernan et al., 2017) をベースに、

(1) 基課題で研究してきた語義曖昧性解消の技術を情報抽出という応用タスクに適用する研究を行う予定であった。また申請者は、

(2) ドメイン適応の技術を本研究に応用し、他分野の論文からの問題点/解決法抽出、

(3) 他言語 (日本語) の論文からの問題点/解決法抽出、また、

(4) レビュー文からの企業のニーズ/問題点解決法の抽出に発展させることを目的としていた。

また、科研費の費目が「国際共同研究強化」であるため、共同研究先のイギリスの Cambridge 大学の Teufel 教授およびその周辺の研究者との関係を深め、日本語と英語の自然言語処理の研究について交流することも大きな目的であったと考えている。

3. 研究の方法

上記「2. 研究の目的」に記述した 4 つの目的について、先行研究と同様に Support Vector Machine を使って順に研究する予定であったが、2018 年の秋に Google 社によって BERT が発表され、2019 年の終わりか 2020 年には、大規模事前学習モデルの利用による自然言語処理が主流となってきた。BERT のような大規模事前学習モデルは大量の事例を必要とせず、少量の事例から Fine-tuning と呼ばれるももとのモデルのパラメータの微細な調整により追加的な学習を行うことで、これまでの最高性能を大幅に上回る結果を出す。そのため、自然言語処理の研究分野の中で、タグ付きコーパスの拡張よりも、事前にタグなしのコーパスで自己学習した大規模事前学習モデルの方が有用であるという知見が得られてきた。

さらに、BERT のような Transformer 型のモデルの出力ベクトルを利用することで、語義曖昧性解消の目的とする、言葉の意味を表現したり分類したりすることがトークンごとに実現可能になり、翻訳などの下段タスクの前処理としての語義曖昧性解消の重要性は下がってきたと考えている。

紆余曲折を経たのち、以下の方法を取ることにした。

(1) まず、対象文書は日本語の自然言語処理の論文とした

(2) Support Vector Machine による BERT を Fine-tuning する方法を利用して、論文からの問題点を含む文の判定システムを作成することとした

(3) 日英の問題の記述の仕方が異なることから、日本語の自然言語処理のコーパスに「問題」を含む文であるかどうかをアノテーションすることとした

(4) 「問題」の語義についても同コーパスにアノテーションを行った

(5) 目的のタスクに、語義曖昧性解消タスクと、文脈の追加が精度向上に有効であるかの調査を行う研究をすることとした

4. 研究成果

申請者は、2018 年の 5 月から 2019 年の 4 月まで、サバティカルを取得して Cambridge 大学に客員研究者として滞在し、大学の研究者と研究交流を行った。渡英中、BERT を発表し

た Google 社(Google London)や、Apple Cambridge、富士通のヨーロッパ研究所、オックスフォード大学を訪問し、日本語の語義曖昧性解消およびドメイン適応について研究発表を行った。

問題抽出については、日本語のコーパスを、『言語処理学会論文誌 LaTeX コーパス』に決定し、英語版のループリックの定義(当時更新中であった)について、出来次第送ってもらった。一方で、語義曖昧性解消の研究と、Fine-tuning を使った言語横断の転移学習の研究を先に行った。問題抽出の論文は現在原著論文を執筆中である。以下の論文や研究会の発表を行った。また、研究交流を通じて、Cambridge 大学との新たな国際共同研究を行う予定である。

著書 1 冊：

- ・柴原 一友，築地 毅，古宮嘉那子，宮武孝尚，小谷 善行，械学習教本，森北出版，共著，(2019, 10)。

原著論文 6 編：

- ・鈴木類，古宮嘉那子，浅原正幸，佐々木稔，新納浩幸，概念辞書の類義語と分散表現を利用した教師なし all-words WSD，自然言語処理，Vol.26, No.2, pp. 361-380, (2019.6)。
- ・Jun Izutsu, Kanako Komiya, Morphological Analyzer Using the Bi-LSTM Model Only for Japanese Hiragana Sentences, International Journal on Natural Language Computing, vol. 11, no. 1, (2022.02)。
- ・Mika Kishino, Kanako Komiya, Extracting Speech Patterns of Japanese Fictional Characters Using Subword Units, International Journal on Natural Language Computing, vol. 11, no. 3, (2022.02)。
- ・Kanako Komiya, Aya Tanabe, Hiroyuki Shinnou, Diachronic Domain Adaptation of Word Sense Disambiguation for Corpus of Historical Japanese Using Word Embeddings, NINJAL Research Papers, vol. 23, pp. 29-57, (2022,07)。
- ・河野慎司，古宮嘉那子，新納浩幸，簡易小型化 BERT による日本語構文解析，自然言語処理，29 巻，3 号，pp. 854-874, (2022,09)。
- ・Kanako Komiya, Shinji Kouno, Takumi Seitou, Teruo Hirabayashi, Composing Word Embeddings for Compound Words Using Linguistic Knowledge, ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), No.: 58, pp 1-22 (2023,03)。

国際会議発表 10 編：

- ・Kanako Komiya, All-words Word Sense Disambiguation in Japanese, 6th Annual Oxbridge Woman In Computer Science Conference 2019, (No proceedings), (2019.3)。
- ・Kanako Komiya, Takumi Seitou, Minoru Sasaki, Hiroyuki Shinnou, Composing Word Vectors for Japanese Compound Words Using Dependency Relations, CICLING 2019, no 229, (2019.4)。
- ・Teruo Hirabayashi, Kanako Komiya, Masayuki Asahara and Hiroyuki Shinnou, Automatic Creation of Correspondence Table of Meaning Tags from Two Dictionaries in One Language Using Bilingual Word Embedding, 13th BUCC Workshop at LREC 2020, pp. 22-28, (2020,5)。
- ・Kanako Komiya, Daiki Yaginuma, Masayuki Asahara, Hiroyuki Shinnou, Generation and Evaluation of Concept Embeddings Via Fine-Tuning Using Automatically Tagged Corpus, PACLIC 2020, pp. 122-128, (2020,10)。
- ・Teruo Hirabayashi, Kanako Komiya, Masayuki Asahara, Hiroyuki Shinnou, Composing Word Vectors for Japanese Compound Words Using Bilingual Word Embeddings, PACLIC 2020, pp. 404-410, (2020,10)。
- ・Masashi Takaku, Toshio Hirasawa, Mamoru Komachi, Kanako Komiya, Neural Machine Translation from Historical Japanese to Contemporary Japanese Using Diachronically Domain-Adapted Word Embeddings, PACLIC 2020, pp. 534-541, (2020,10)。
- ・Jun Izutsu, Kanako Komiya, Morphological Analysis of Japanese Hiragana Sentences Using the Bi-LSTM CRF Model, 10th International Conference on Natural Language Processing (NLP 2021), (2021, 12, 23)。
- ・Mika Kishino, Kanako Komiya, Extraction of Linguistic Speech Patterns of Japanese Fictional Characters Using Subword Units, 10th International Conference on Natural Language Processing (NLP 2021), (2021, 12, 23)。
- ・Kanako Komiya, Nagi Oki and Masayuki Asahara, Word Sense Disambiguation of Corpus of Historical Japanese Using Japanese BERT Trained with Contemporary Texts, PACLIC 2022, (2022, 10, 20)。
- ・Jun Izutsu, Kanako Komiya, and Hiroyuki Shinnou, Word Segmentation of Hiragana

Sentences Using Hiragana BERT, PRICAI 2023, (2023,11,17).

・ Shoma Asada, Kanako Komiya, and Masayuki Asahara, All-Words Word Sense Disambiguation for Historical Japanese, PACLIC 2023, (2023,12,03).

研究会 20 編：

・ 柳沼 大輝, 古宮 嘉那子, 新納 浩幸, All-words WSD と fine-tuning を利用した分類語彙表の語義の分散表現の構築, 言語資源活用ワークショップ 2019, pp. 179-184, (2019.09.02).

・ 佐々木 稔, 古宮嘉那子, 単語区切りの違いによる QA サイトの質問回答ペアの分類, IDR ユーザフォーラム 2019, (2019.11.29). (*ヤフー賞受賞)

・ 井筒順, 明石陸, 加藤涼, 岸野望叶, 小林汰一郎, 金野佑太, 古宮嘉那子, MeCab による平仮名のみ形態素解析, 言語処理学会第 26 回年次大会, pp. 65-69, (2020,03,17).

・ 河野慎司, 古宮嘉那子, マルチタスク学習を利用した短単位の分散表現から長単位の分散表現の合成, 言語処理学会第 26 回年次大会, pp. 209-212, (2020,03,17).

・ 高久雅史, 平澤寅庄, 小町守, 古宮嘉那子, 通時的な領域適応を行った単語分散表現を利用した古文から現代文へのニューラル機械翻訳, 訂正原稿, 言語処理学会第 26 回年次大会, pp. 677-680, (2020,03,18).

・ 平林照雄, 古宮嘉那子, 新納浩幸, Bilingual Word Embeddings による短単位と長単位のアラインメント, 言語処理学会第 26 回年次大会, pp. 1257-1260, (2020,03,19).

・ 佐々木 稔, 古宮嘉那子, 複数の事前学習済みモデルを用いた QA サイト質問回答ペアの分類, IDR ユーザフォーラム 2020, (2020.11.24).

・ 河野 慎司, 古宮嘉那子, 品詞情報を利用した複合語の分散表現の合成, 音声言語および自然言語処理シンポジウム, 研究報告自然言語処理 (NLP), Vol.2020-NL-246, No.23, pp.1-6, (2020.12.03).

・ 井筒順, 古宮嘉那子, Bi-LSTM CRF モデルを用いた平仮名文の形態素解析, 言語処理学会第 27 回年次大会, pp. 942-946, (2021,03,17).

・ 三戸尚樹, 古宮嘉那子, 佐々木 稔, 共学習によるレビュー文書からのネガティブな意見文の抽出, 言語処理学会第 28 回年次大会, pp. 1253-1257, (2022,03,16).

・ 河野 稜斗, 平林照雄, 古宮嘉那子, BERT を用いた二つの辞書の対応付け, 言語処理学会第 28 回年次大会, pp. 746-751, (2022,03,16).

・ 多喜 凧, 古宮嘉那子, 現代文 BERT を利用した日本語歴史コーパスの語義曖昧性解消, 訂正原稿, 言語処理学会第 28 回年次大会, pp. 1931-1935, (2022,03,17).

・ 平林照雄, 古宮嘉那子, 浅原正幸, 科学技術論文における「問題」の周辺文の問題内容の抽出, 言語資源ワークショップ 2022, (2022,08,30).

・ 岸野望叶, 古宮嘉那子, 新納浩幸, T5 による特定キャラクター風発話への変換とその言語モデルの構築, 第 253 回自然言語処理研究発表会, (2022, 09, 30).

・ 浅田宗磨, 古宮嘉那子, 日本語歴史コーパスの All-words WSD, 言語処理学会第 29 回年次大会, (2023, 03, 14).

・ 臼井久生, 古宮嘉那子, T5 を用いた古文から現代文への翻訳, 言語処理学会第 29 回年次大会, (2023, 03, 16).

・ 古宮嘉那子, 日本語学会 2023 年度春季大会 シンポジウム「情報技術と大規模テキスト資源がひらく日本語史研究」にて「近代以前の日本語を対象にした自然言語処理の紹介」, (2023, 05, 21)

・ 古宮嘉那子, 語彙・辞書研究会 秋の研究会にて「語義曖昧性解消 コーパスへの意味タグの付与システム」, (2023, 11, 18)

・ Ganbat Naranbuvei, 浅田宗磨, 古宮嘉那子, 翻訳と BabelNet を利用した日本語の語義曖昧性解消, 言語処理学会第 30 回年次大会, (2024, 03, 12)

・ 浅田 宗磨, 古宮嘉那子, 浅原 正幸, 『現代日本語書き言葉均衡コーパス』に対する分類語彙表番号悉皆付与, 言語処理学会第 30 回年次大会, (2024, 03, 14)

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 3件）

1. 著者名 Komiya Kanako, Kono Shinji, Seito Takumi, Hirabayashi Teruo	4. 巻 22
2. 論文標題 Composing Word Embeddings for Compound Words Using Linguistic Knowledge	5. 発行年 2023年
3. 雑誌名 ACM Transactions on Asian and Low-Resource Language Information Processing	6. 最初と最後の頁 1~22
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3561299	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Shinji Kono, Komiya Kanako, Hiroyuki Shinnou	4. 巻 29
2. 論文標題 Japanese Parsing Using Smaller BERT	5. 発行年 2022年
3. 雑誌名 Journal of Natural Language Processing	6. 最初と最後の頁 854~874
掲載論文のDOI（デジタルオブジェクト識別子） 10.5715/jnlp.29.854	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 古宮 嘉那子、田邊 絢、新納 浩幸	4. 巻 23
2. 論文標題 分散表現を利用した日本語歴史コーパスにおける語義曖昧性解消の通時適応	5. 発行年 2022年
3. 雑誌名 国立国語研究所論集 = NINJAL Research Papers	6. 最初と最後の頁 59~73
掲載論文のDOI（デジタルオブジェクト識別子） 10.15084/00003566	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Jun Izutsu, Kanako Komiya	4. 巻 11
2. 論文標題 Morphological Analyzer Using the Bi-LSTM Model Only for Japanese Hiragana Sentences	5. 発行年 2022年
3. 雑誌名 International Journal on Natural Language Computing	6. 最初と最後の頁 29-45
掲載論文のDOI（デジタルオブジェクト識別子） 10.5121/ijnlc.2022.11103	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Mika Kishino, Kanako Komiya	4. 巻 11
2. 論文標題 Extracting Speech Patterns of Japanese Fictional Characters Using Subword Units	5. 発行年 2022年
3. 雑誌名 International Journal on Natural Language Computing	6. 最初と最後の頁 1-14
掲載論文のDOI (デジタルオブジェクト識別子) 10.5121/ijnlc.2022.11101	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Suzuki Rui, Komiya Kanako, Asahara Masayuki, Sasaki Minoru, Shinnou Hiroyuki	4. 巻 26
2. 論文標題 Unsupervised All-words WSD Using Synonyms and Embeddings	5. 発行年 2019年
3. 雑誌名 Journal of Natural Language Processing	6. 最初と最後の頁 361 ~ 379
掲載論文のDOI (デジタルオブジェクト識別子) 10.5715/jnlp.26.361	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計30件 (うち招待講演 0件 / うち国際学会 6件)

1. 発表者名 Kanako Komiya, Nagi Oki and Masayuki Asahara
2. 発表標題 Word Sense Disambiguation of Corpus of Historical Japanese Using Japanese BERT Trained with Contemporary Texts
3. 学会等名 The 36th Pacific Asia Conference on Language, Information and Computation (国際学会)
4. 発表年 2022年

1. 発表者名 臼井久生, 古宮嘉那子
2. 発表標題 T5を用いた古文から現代文への翻訳
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2022年

1. 発表者名 浅田宗磨, 古宮嘉那子
2. 発表標題 日本語歴史コーパスのAll-words WSD
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2022年

1. 発表者名 岸野望叶, 古宮嘉那子, 新納浩幸
2. 発表標題 T5 による特定キャラクター風発話への変換とその言語モデルの構築
3. 学会等名 第253回自然言語処理研究発表会
4. 発表年 2022年

1. 発表者名 平林照雄, 古宮嘉那子, 浅原正幸
2. 発表標題 科学技術論文における「問題」の周辺文からの問題内容の抽出
3. 学会等名 言語資源ワークショップ2022
4. 発表年 2022年

1. 発表者名 Jun Izutsu, Kanako Komiya
2. 発表標題 Morphological Analysis of Japanese Hiragana Sentences Using the Bi-LSTM CRF Model
3. 学会等名 10th International Conference on Natural Language Processing (NLP 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Mika Kishino, Kanako Komiya
2. 発表標題 Extraction of Linguistic Speech Patterns of Japanese Fictional Characters Using Subword Units
3. 学会等名 10th International Conference on Natural Language Processing (NLP 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 多喜 凧, 古宮嘉那子
2. 発表標題 現代文 BERT を利用した日本語歴史コーパスの語義曖昧性解消
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2021年

1. 発表者名 河野稜斗, 平林照雄, 古宮嘉那子
2. 発表標題 BERTを用いた二つの辞書の対応付け
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2021年

1. 発表者名 三戸尚樹, 古宮嘉那子, 佐々木稔
2. 発表標題 共学習によるレビュー文書からのネガティブな意見文の抽出
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2021年

1. 発表者名 HUANG YIPU, 佐々木稔, 古宮嘉那子
2. 発表標題 レビューから抽出されたキーフレーズと感情スコアを用いた評判分析
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2021年

1. 発表者名 Teruo Hirabayashi, Kanako Komiya, Masayuki Asahara and Hiroyuki Shinnou
2. 発表標題 Automatic Creation of Correspondence Table of Meaning Tags from Two Dictionaries in One Language Using Bilingual Word Embedding
3. 学会等名 13th BUCC Workshop at LREC 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Kanakano Komiya, Daiki Yaginuma, Masayuki Asahara, Hiroyuki Shinnou
2. 発表標題 Generation and Evaluation of Concept Embeddings Via Fine-Tuning Using Automatically Tagged Corpus
3. 学会等名 PACLIC 2020
4. 発表年 2020年

1. 発表者名 Teruo Hirabayashi, Kanako Komiya, Masayuki Asahara
2. 発表標題 Hiroyuki Shinnou, Composing Word Vectors for Japanese Compound Words Using Bilingual Word Embeddings
3. 学会等名 PACLIC 2020
4. 発表年 2020年

1. 発表者名 Masashi Takaku, Toshi Hirasawa, Mamoru Komachi, Kanako Komiya
2. 発表標題 Neural Machine Translation from Historical Japanese to Contemporary Japanese Using Diachronically Domain-Adapted Word Embeddings
3. 学会等名 PACLIC 2020
4. 発表年 2020年

1. 発表者名 佐々木稔, 古宮嘉那子
2. 発表標題 複数の事前学習済みモデルを用いたQAサイト質問回答ペアの分類
3. 学会等名 IDRユーザフォーラム 2020
4. 発表年 2020年

1. 発表者名 河野 慎司, 古宮嘉那子
2. 発表標題 品詞情報を利用した複合語の分散表現の合成
3. 学会等名 音声言語および自然言語処理シンポジウム
4. 発表年 2021年

1. 発表者名 井筒順, 古宮嘉那子
2. 発表標題 Bi-LSTM CRF モデルを用いた平仮名文の形態素解析
3. 学会等名 言語処理学会第27回年次大会
4. 発表年 2021年

1. 発表者名 平林照雄, 河野慎司, 古宮嘉那子, 新納浩幸
2. 発表標題 日本語の論文コーパスにおける「問題」の語義アノテーション
3. 学会等名 言語処理学会第27回年次大会
4. 発表年 2021年

1. 発表者名 岸野望叶, 古宮嘉那子
2. 発表標題 SentencePieceを用いたキャラクターの特徴語抽出
3. 学会等名 言語処理学会第27回年次大会
4. 発表年 2021年

1. 発表者名 金野佑太, 古宮嘉那子
2. 発表標題 論文の要旨からのタイトル生成におけるキーワードと分野別fine-tuningの効果
3. 学会等名 言語処理学会第27回年次大会
4. 発表年 2021年

1. 発表者名 小林汰一郎, 古宮嘉那子
2. 発表標題 SVMを用いたBCCWJにおける同形異音語の読み推定
3. 学会等名 言語処理学会第27回年次大会
4. 発表年 2021年

1. 発表者名 Kanao Komiya, Takumi Seitou, Minoru Sasaki, Hiroyuki Shinnou
2. 発表標題 Composing Word Vectors for Japanese Compound Words Using Dependency Relations
3. 学会等名 CICLING 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 柳沼 大輝, 古宮 嘉那子, 新納 浩幸
2. 発表標題 All-words WSDとfine-tuningを利用した分類語彙表の語義の分散表現の構築
3. 学会等名 言語資源活用ワークショップ 2019
4. 発表年 2019年

1. 発表者名 佐々木稔, 古宮嘉那子
2. 発表標題 単語区切りの違いによるQAサイトの質問回答ペアの分類
3. 学会等名 IDRユーザフォーラム 2019
4. 発表年 2019年

1. 発表者名 井筒順, 明石陸, 加藤涼, 岸野望叶, 小林汰一郎, 金野佑太, 古宮嘉那子
2. 発表標題 MeCab による平仮名だけの形態素解析
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 河野慎司, 古宮嘉那子
2. 発表標題 マルチタスク学習を利用した短単位の分散表現から長単位の分散表現の合成
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 高久雅史, 平澤寅庄, 小町守, 古宮嘉那子
2. 発表標題 通時的な領域適応を行った単語分散表現を利用した古文から現代文へのニューラル機械翻訳
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 平林照雄, 古宮嘉那子, 新納浩幸
2. 発表標題 Bilingual Word Embeddingsによる短単位と長単位のアラインメント
3. 学会等名 語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 Kanao Komiya
2. 発表標題 All-words Word Sense Disambiguation in Japanese
3. 学会等名 6th Annual Oxbridge Woman In Computer Science Conference 2019 (国際学会)
4. 発表年 2019年

〔図書〕 計1件

1. 著者名 柴原 一友、築地 毅、古宮 嘉那子、宮武孝尚、小谷 善行	4. 発行年 2019年
2. 出版社 森北出版	5. 総ページ数 240
3. 書名 機械学習教本	

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
トイフェル シモーネ (Teufel Simone)	ケンブリッジ大学・Department of Computer Science and Technology・Professor	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関		
イギリス	University of Cambridge		