

論文内容の要旨

博士論文題目 Automatic Error Tag Annotation on the Writing of Japanese Language Learners for Linguistic and Educational Research
(日本語学習者の作文コーパスの言語教育研究のための誤用タグアノテーションの自動化)

氏名 大山 浩美

(論文内容の要旨)

近年、様々な種類の言語学習者コーパスが収集され、言語教育の調査研究に利用されている。ウェブを利用した言語学習アプリケーションも登場し、膨大な量のコーパスを収集することも可能になってきている。学習者が生み出した文には正用だけでなく誤用も含まれており、それらの大規模な誤用文を言語学や教育などの研究に活かすことが重要である。日本語教育の現場では、学習者の書いた作文において学習者が誤りを犯す原因を追及し、誤用を犯さないようにフィードバックとして活かしたい必要があるが、大規模な言語学習者コーパスを調査分析するのは困難である。研究に活かすために、学習者コーパス内の学習者の書いた誤用を検出し、それらに誤用の種類を明示する誤用タグを振り、統計的に分析することが重要である。そのような理由から、本研究では以下の作業、実験を行い、次のようなことが分かった。

1. 日本語学習者の作文に誤用タグを付与した誤用コーパス (NAIST 誤用コーパス) を作成した。(a) 関連する日本語学習者コーパスで使用されている誤用タグについて調査を行った。(b) 現存する誤用タグの長所短所を考慮し、汎用性があるような誤用タグを作成した。(c) NAIST 誤用コーパスを用い、日本語学習者の格助詞の誤用について調査した。その助詞誤用頻度分析の結果、助詞を脱落させる誤りが最も多いことがわかった。さらに、助詞「の」や「は」において学習者の習得の難しさが見られた。
2. NAIST 誤用コーパスにおいて機械学習法を用いた誤用タイプ別自動分類実験を行い、かつアプリケーションに堪えうる適合率を実現した(8割程度)。(a) 誤用タイプ別自動分類実験をドメイン外のコーパスにおいても行ったが、NAIST 誤用コーパスでの実験に比べると精度が14.9%ほど低かった。
3. 新聞コーパスにおいて格助詞「を」の正用例、誤用例抽出実験を行った。(a) 「を」の正用モデルを新聞コーパスから作り、そのモデルを用い、学習者のコーパスで誤用例を判定できるかどうかの実験を行った。100事例の場合、F値で50.0%、200事例の場合、F値で53.9%の精度で判定できた

以上の作業を行い、学習者コーパスの整備をすることにより、言語教育の調査研究に関して新しい知見が得られた。

氏名	大山 浩美
----	-------

(論文審査結果の要旨)

平成 27 年 12 月 22 日に開催した公聴会の結果を参考に平成 28 年 2 月 22 日に本博士論文の審査を行った。以下のとおり、本博士論文は、提案者が独立した研究者として、研究活動を続けていくための十分な素養を備えていることを示すものと認める。

大山 浩美は、本博士論文において、日本語学習者の作文中の誤りのタイプに関する詳細な調査を行い、誤りの種類を分類することによって、新しい誤りタイプのセットを提案した。また、その分類に従って、学習者の日本語作文の誤り訂正データに対して誤りタイプの自動分類を行う手法を提案し、評価を行った。さらに、学習者の誤りの中でも頻度の高い助詞誤りを取り上げ、自動誤り訂正手法を提案し、実験による精度評価を行った。本論文は、日本語学習者の作文中の誤り訂正と分類に関して、次のような貢献を認める。

1. 日本語学習者の誤りの詳細な解析を行い、言語学的にも自然であり、かつ、計算機による分類にも適切な誤りタイプの分類を提案した。
2. 日本語学習者の作文中の誤りに対して誤用タグを付与したコーパス (NAIST 誤用コーパス) を構築し、研究コミュニティに公開した。
3. 構築した NAIST 誤用コーパスに対して、誤用タイプの自動分類実験を行い、その性能評価を行った。また、自動分類アプリケーションとして利用可能な程度の分類性能を報告した。
4. 新聞コーパスを利用して、日本語の助詞の使用誤りに対する自動誤り訂正法を提案し、実験によって精度票を行った。特に、格助詞「を」の正用例、誤用例の分類実験を行い、問題点や改良点について考察を行った。

日本語学習者の文法誤りの種類に対する新たな誤り分類を提案し、実データに対する自動的な誤りタイプ分類法の提案と評価を行った本研究は、実用的であり、自然言語処理技術を利用した言語教育分野において高い貢献があると評価する。

よって、本論文は、博士（工学）の学位論文として価値あるものと認める。