

研究タイトル：

文法誤り訂正と第二言語習得支援



氏名： 小田 幹雄 / ODA Mikio E-mail: oda@kurume-nct.ac.jp

職名： 教授 学位： 博士（工学）

所属学会・協会： 言語処理学会，情報処理学会，日本知能情報ファジィ学会

キーワード： 自然言語処理，第二言語習得

 技術相談：

- ・自然言語処理一般

 提供可能技術：

- ・第二言語習得支援

- ・機械学習

研究内容： 自然言語処理を用いた文法誤り訂正と第二言語習得支援システムの開発

自然言語処理は、大量のデータと機械学習の発展により、文書データの分析と検索、文章の予測、言語翻訳、質問応答、感情分析など、さまざまな応用例に活用できる。その中で、作成した文章の文法の誤りを訂正するタスクである文法誤り訂正（Grammatical Error Correction）は、非言語ネイティブの第二言語（L2）学習者や学習機械にとって、困難な課題であり、特に、データ駆動型のGEC学習には、訓練された言語ネイティブの人間により作成された大規模の訓練データが必要である。しかしながら、より大規模の訓練データを作成することは困難であるため、人工的に合成された大規模な訓練データも併用され、このとき、GECシステムの学習に関して、その人工合成訓練データの品質が、大きな問題となる。

本研究では、人工合成訓練データに関して、人工的な間違いを挿入する前の原文ではなくGECモデルの予測文が、より適切な誤り訂正文であるという仮説を立てて、人工合成訓練データを再構築する方法を提案する。実験の結果、提案する再構築された人工合成訓練データによる学習は、元の人工合成訓練データによる学習よりも優れていることが示された。また、本提案手法による人工学習訓練データのみを用いた学習は、L2学習者の誤り文章から生成された学習訓練データと従来の人工合成訓練データを併用した学習と同等の効果が得られることが示された。

さらに、L2学習者による文章作成支援システムを開発する場合、文法誤り訂正機能を有することが有効であり、本研究の成果による推論システムの組み込みが可能である。また、文章作成支援システムにおいては、ユーザへの抽象意味表現による提案も有効であると考えられる。現在、抽象意味表現から文章への変換やその逆変換の精度は、文法誤り訂正の精度より低く、より精度のよい抽象意味表現変換モデルが望まれている。本研究による仮説は、抽象意味表現の学習にも応用可能であり、統語論と意味論の両面からの文章作成支援の開発に発展できると考える。

提供可能な設備・機器：

名称・型番（メーカー）

GPU・RTX A6000(NVIDIA)