

論文 / 著書情報  
Article / Book Information

題目(和文)	
Title(English)	Parsing Argumentative Structure in English-as-a-Foreign-Language Learner Essays
著者(和文)	Jan Wira Gotama Putra
Author(English)	Jan Wira Gotama Putra
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12127号, 授与年月日:2021年9月24日, 学位の種別:課程博士, 審査員:徳永 健伸,岡崎 直観,村田 剛志,宮崎 純,齋藤 豪
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12127号, Conferred date:2021/9/24, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

## 論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	Jan Wira Gotama PUTRA	
論文審査 審査員		氏名	職名	氏名	職名
	主査	徳永健伸	教授	斎藤豪	准教授
	審査員	岡崎直観	教授		
		村田剛志	教授		
宮崎純		教授			

### 論文審査の要旨 (2000 字程度)

本論文は"Parsing Argumentative Structure in English-as-a-foreign-language Learner Essays"と題し、英文6章と付録2章から構成されており、英語学習者が書いたエッセイの議論構造を解析する手法を提案している。また、その応用として、解析結果の構造に基づいて文順序を並べ替えてテキストを改善する手法を提案している。

第1章"Introduction"では、本論文の背景と目的について述べている。議論構造解析はテキストの生成、評価、要約、改善などのさまざまな問題を解決するための基本的な要素技術である。本研究では言語学習支援への応用を前提とし、従来の議論構造解析が対象としてこなかった言語学習者が書いたテキストを対象とすると述べている。また、解析結果を利用し、学習者のテキストを改善するための指針を与えることは言語学習において大きな意義があると述べている。

第2章"Related Work"では、議論構造マイニング、議論構造研究のためのコーパス、議論構造解析、文集合の整列課題の4つの関連分野について先行研究を概観し、本研究の位置付けおよび貢献について明らかにしている。

第3章"Corpus Construction and Annotation Study"では、本研究を進める上で不可欠となるコーパスの構築について述べている。本章は、(1)コーパスに付与する情報の設計とその手順、(2)アノテーションを効率的かつ正確におこなうためのアノテーション・ツールの開発、(3)テキストの構造を考慮した新しいアノテーションの信頼性の評価尺度の提案、(4)アノテーション結果の信頼性・品質に関する分析と議論の4つのトピックを含んでいる。本研究で構築されたコーパス ICNALE-AS2R は、神戸大学を中心に収集されたアジアの大学生が書いたエッセイコーパス ICNALE の一部(434 エッセイ)に議論構造とエッセイ改善のアノテーションをおこなったもので、改善の前後を含む言語学習者のコーパスとして世界初のものであると述べている。

第4章"Argumentative Structure Parsing"では、まず議論構造解析に関連する文対を同定するリンキング課題と関連する文対の関係を分類するラベリング課題に分け、それぞれについて定式化をおこない、ICNALE-AS2R コーパスを用いて既存の最高性能モデルを使った実験について述べている。複数のニューラルネットワークモデルを用いて実験をおこなった結果、リンキング課題については Sentence-BERT で文をエンコードし、Biaffine モデルで関係付けをおこなうモデルがもっともよい結果を示し、ラベリング課題では BERT によるファインチューニングモデルがもっともよい結果を示したと報告している。これらをベースラインとして、本研究では特にリンキング課題に注目し、議論構造における文の位置を予測する副タスクをリンキングに加えたマルチタスク学習を導入し、性能を向上できたと報告している。さらに訓練データの量を増やすために、ICNALE-AS2R コーパスに加え、議論構造を付与した既存のエッセイコーパス (PEC) を選択的に用いてモデルを学習することによって F1 マクロ値を 15.8%改善できたと報告している。

第5章"Automatic Sentence Reordering"では、言語学習者が書いたテキストを文の順序を並び替えることによって改善する課題を提案し、議論構造の解析結果を使って文順序を並び替える手法を提案している。従来から文の順序を整列する課題については研究されていたが、従来は文の集合を整列する問題であったのに対し、本研究が対象とするのは既にテキストとしての体裁を持った文の並びを一部並び替えてテキスト全体の構成をさらによくする課題であり、より困難な課題となっていると述べている。提案手法は、(1) 談話構造解析、(2) 談話構造解析で関係付けられた文対の局所的順序の決定、(3) 文の全順序の決定の3つの段階から構成される。ICNALE-AS2R を用いた実験をおこない、言語生成で提案されている Maximum Local Coherence モデル、文の並び替え課題で最高性能を報告している Topological Sorting モデルと比較した結果、提案手法が著しく性能が高かったと報告している。しかしながら、提案手法の出力は改善前の元テキストを越えることはできなかったことから、エラー分

析をおこなった結果、議論構造解析の精度が隘路となっていることを明らかにしている。第7章"Conclusions"では、本論文の貢献が以下の3点であることを述べている：(1) エッセイの改善結果も含む言語学習者のエッセイコーパスの構築と評価、(2) 言語学習者の書いたエッセイの議論構造解析手法の提案、(3) テキストの改善のための文の並べ替え課題の提案とそのための手法の提案。以上要するに、本論文は言語学習者が書いた不完全なエッセイを対象とし、コーパスの構築、議論構造解析手法の提案、文の並べ替えによるエッセイの改善手法の提案をおこない、その有効性を評価実験により示すとともに、議論構造をテキストの改善に応用することの有用性を示している。本論文の成果はコンピュータによる言語学習支援のための基盤技術として位置付けることができ、工業上貢献するところが大きい。よって、本論文は博士(工学)の学位論文として十分価値あるものと認める。

注意：「論文審査の要旨及び審査員」は、東工大リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。