

令和 3 年 6 月 7 日現在

機関番号：12501
研究種目：若手研究(B)
研究期間：2016～2020
課題番号：16K16885
研究課題名(和文)ライティングプロセス可視化コーパスの構築と英語学習者のライティングプロセスの解明

研究課題名(英文)Development of writing process corpus and elucidation of English learners' writing process

研究代表者
石井 雄隆(Ishii, Yutaka)

千葉大学・教育学部・准教授

研究者番号：90756545
交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：本研究の目的は、コンピュータのキー入力記録を用いたライティングプロセス可視化コーパスの構築と英語学習者のライティングプロセスの解明である。はじめに、キー入力記録システムを用いて、日本人英語学習者のライティングプロセスデータを収集し、英語学習者のキー入力記録情報を含んだ学習者コーパスを構築した。その後、それらのデータにライティングの評価や品詞情報などのアノテーションを行い、最後に、完成したプロダクトに関する指標とライティング執筆中のプロセスの指標を用いて、プロセスとプロダクトの関係性や評価に寄与する特徴量などを調査し、英語学習者のキー入力ログを用いた新しいライティング研究の可能性を検討した。

研究成果の学術的意義や社会的意義

本研究課題の意義は、英語学習者のライティングプロセスという学習者のログや学習履歴を活用した新たなライティング研究や、データマイニングを活用した新たな外国語教育研究、学習者のライティング能力の発達を記述するための新しい教育評価研究に活用することが可能である。また、本研究の成果はライティング研究のみではなく、学習者コーパス研究さらには、第二言語習得研究にも貢献が可能である。

研究成果の概要(英文)：This study aims to develop English learners' writing process corpus and elucidate English learners' writing process. First of all, this project collected learners' writing process data and developed learner corpus. After that, the data was annotated. Lastly, the relationship between learners' writing product and process was investigated.

研究分野：外国語教育

キーワード：英語教育 第二言語ライティング キー入力記録 学習者コーパス ICT 測定・評価

様式 C - 19 , F - 19 - 1 , Z - 19 (共通)

1. 研究開始当初の背景

近年、大量の学習者の情報に基づいて教育における意思決定を行うために、エデュケーションナルデータマイニングやラーニングアナリティクスと呼ばれる分野が注目を受けており、英文読解時の眼球運動の測定や、コンピュータ上で問題を回答した際のマウスの軌跡などを含んだ学習ログの分析など英語学習者の学習プロセスに関するデータの収集・分析とそれに基づく教育改善が注目を集めている。

日本における英語教育では、グローバル化が進行する国内外の情勢を受け、高度な英語力を持つ人材の育成が強調されている。それに伴い、国際社会で活躍できる人材養成の社会的要請にこたえるため、国内の大学において、ライティングセンターの設置などが進んでおり、ライティングとその指導に関する関心は高まってきている。

英語ライティング研究においては、刺激再生法や思考発話法などの内省的研究法を用いたプロセス研究はこれまでも行われてきた。しかしながら、学習者の行動履歴であるライティングのキー入力ログを活用したライティングプロセスに関する研究は、これまでにあまり行われていない。本研究課題は、そこに焦点を当て、コンピュータを使うことでしか知ることができない英語学習者のキー入力ログに基づいたライティングプロセスの解明を目指す。これまで申請者が行ってきたライティング研究は、完成したプロダクトを主な対象としてきたが、それは学習者のライティングの一部を表しているに過ぎない。学習者は執筆中も、計画・文章化・推敲などの複雑な認知プロセスをたどっている。それらを用いた学習者のライティングプロセスの検討は、コンピュータによる言語テストの普及の議論を鑑みると、今後のライティング研究に大きな示唆を与えることが可能である。

本申請課題においては、コンピュータを用いてしか知ることができない英語学習者のライティングプロセスのモデル化やライティングプロダクトに影響を与えるライティングプロセス指標の検討、またそれらの結果に基づく新しいライティング指導と評価の可能性を検討する。

2. 研究の目的

初めに、英語学習者のライティングプロセスについて検討するために、大規模な学習者のデータを収集し、コーパスを構築する。収集したデータを用いて、ライティングプロダクト及びプロセスに関する情報をデータに付与する。プロダクトに関しては、L2 Syntactic Complexity Analyzer (Lu, 2010) などを用いて、作文データに文の複雑さ指標や品詞情報を付与し、また文法的誤り情報などの付与も検討する。データへのアノテーションを終えたら、下記 4 つの研究課題について明らかにする。

(1) 学習者のプロセスとプロダクトの相互作用

プロセスに関する指標であるポーズや推敲などがプロダクトにおける英文の複雑さ・正確さ・流暢さにどのように関係しているかや、またライティング執筆時の推敲過程はプロダクトにおける文の正確さとどのように関係するかなどを分析する。

(2) 英語学習者のライティングプロセスのモデル化

時系列データ分析 (time series data analysis) の手法を用いることによって、個々人のライティングプロセスをより詳細に捉え、ライティングプロセスのモデル化を試みる。

(3) ライティング執筆時の認知プロセスの傾向

Yamanishi (2009) によって開発されたライティングストラテジーに関する質問紙や Ishikawa (2011) によって開発されたタスクの複雑さの主観的評価に関する質問紙を用いて、実際のキー入力ログと質問紙調査の結果の関係性や、習熟度別の学習者のライティングプロセスの傾向などを調査する。

(4) ライティング自動評価のためのプロセス指標の基礎的検討

複雑さ・正確さ・流暢さなどのプロダクトに関する指標だけではなく、ポーズや推敲などのプロセスに関する変数を用いたライティング自動評価の可能性を検討する。

上記の結果を踏まえ、英語学習者のライティングプロセス及びプロダクトとの関係性を調査し、新しいライティング指導や評価のあり方について検討する。

3. 研究の方法

初めに大規模な学習者のデータを収集し、コーパスを構築する。学習者コーパスの構築に当たっては、李・石川・砂川 (2012) によると、下記 7 つの作業が必要であると指摘されている。

(1) コーパスデザイン作成

(2) 著作権処理用の書類作成

(3) データ収集作業の実行

- (4) 電子化作業の実行
- (5) 誤用タグの検討及び挿入
- (6) 検索エンジンの検討および開発
- (7) データ共有の方法の検討

本研究計画もこれに基づき、データ収集を行う前にコーパスのデザインなどを詳細に検討する。その後、著作権処理用の書類を作成し、コーパス構築の要となるデータの収集を行う。コーパス構築の際に検討しなければならない点として、投野（2013）は、目標言語（モード、ジャンル、文体、トピック）、タスク（データ採取、誘出、参考図書、時間制限）、学習者（認知的・情意的側面、母語背景、L2 学習環境、L2 習熟度）について挙げている。本コーパスは、“It is important for college students to have a part time job.”と“Smoking should be completely banned at all the restaurants in the country.”という二つのトピックを基にデータを収集する。これらは、Ishikawa(2013)によって収集されているアジア人英語学習者の大規模コーパス ICNALE のトピックに準拠している。また、収集に際し、参考図書の閲覧は不可とし、時間制限も ICNALE 準拠で 20 分から 40 分間とする。最後に学習者の情報としては、性別、年齢、大学名・専攻・学年、視覚（英語テストのスコア）の取得状況、英語学習歴、海外滞在歴、英語の使用頻度（5 段階評価）、作文を書くことに対する自信度（5 段階評価）に加えて、Ishikawa（2011）によって開発された concentration, time pressure, anxiety, stress, difficulty, interest, ability, motivation から構成されるタスク遂行に関する主観的困難度や Yamanishi（2009）によって開発された Global Planning (Passage Level), Local Planning (Word/Phrase/Sentence Level), Review/Revision, Avoidance の 4 つの観点から構成されるライティングプロセスに関する質問紙を用いる。データを収集するために、本研究で用いるのは、キー入力記録システム WritingMaetriX である（草薙・阿部・福田・川口、2015）。このシステムでは、記録モードでは入力のためのテキストボックスが表示され、そこに文字が入力される時に入力された文字と入力された時間が記録される。また再生モードでは記録モードで出力されたキーログファイルを使い、書き始めから終了時までの文字入力を動画のように再生することができる。分析モードでは、(a) 総語数、(b) 初入力時間、(c) 一分あたりのキー入力数、(d) 一分あたりの語数、(e) 削除キーの打鍵数に基づいた一分あたりの推敲回数、(f) 前半 / 中盤 / 後半の推敲回数などを確認することが可能である。本研究課題は、このシステムを用いて英語学習者のライティングプロセスデータを収集する。

具体的なデータ収集法としては、国内の私立大学・国公立大学において、データ収集を依頼し、キー入力記録システムである WritingMaetriX を用いて、ライティングデータを収集する。学習者及び著作権処理に関わる情報を収集する際には、オンラインフォームを活用する。

次に、データへのアノテーション及び電子化を開始する。L2 Syntactic Complexity Analyzer (Lu, 2010) を用いて、作文データに文の複雑さ指標の付与、CLAWS を用いた品詞タグの付与、ライティング評価の付与、また和泉・内元・井佐原（2004）によって開発された NICT JLE (Japanese Learner English) Corpus に基づく 47 種類のエラータグの付与を計画している。しかしながら、NICT JLE Corpus は話し言葉のデータなので、本コーパス構築においては、書き言葉用にアレンジする。品詞タグやエラータグのみではなく XML フォーマットによる形式変換、またプロセスに関しては、推敲やポーズなどのタグ付与を検討しているが、実際のデータに合わせて、どのようなアノテーションが適切であるかを検討する。

最後に、収集し、アノテーションを付与したコーパスに対して、様々な指標を用いて、実験を行う。具体的には、ライティングプロダクトの言語的特徴（複雑さ・正確さ・流暢さ）とプロセスの推敲パターンやポーズなどの関連性や、時系列分析の手法を用いた学習者のライティングプロセスのモデル化、ライティング執筆時の認知プロセスの傾向などを検討する。

4. 研究成果

収集した日本人英語学習者のライティングプロセスとライティングプロダクトのデータにアノテーションを行い、様々な分析を行った。学習者のプロセスとプロダクトの相互作用については、プロセスに関する指標であるポーズや推敲などがプロダクトにおける英文の複雑さ・正確さ・流暢さにどのように関係しているかについて分析を行い、日本人英語学習者固有の要因について明らかにした。また英語学習者のライティングプロセスのモデル化については、時系列データ分析の手法を用いることによって、思考発話法や刺激再生法などの手法によって導き出されたライティングプロセスのモデルを時系列データによって再現性を確認した。次に、ライティング執筆時の認知プロセスの傾向については、刺激再生法とキー入力記録を用いた分析を行った。その結果、ライティング執筆時の思考過程について明らかとなり、プロトコルデータを用いた妥当性の検証などについての知見も得られた。最後に、ライティング自動評価のためのプロセス指標の基礎的検討として、『英語教育における自動採点：現状と課題』という書籍を出版した。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 0件）

1. 著者名 石井雄隆・近藤悠介	4. 巻 1
2. 論文標題 ラーニング・アナリティクスと英語ライティング研究	5. 発行年 2017年
3. 雑誌名 JACET-ICT調査研究特別委員会最終報告書	6. 最初と最後の頁 21-28
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
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 石井雄隆	4. 巻 65(3)
2. 論文標題 英語学習者のライティング・プロセスの解明 キー入力記録を用いたプロセスの可視化	5. 発行年 2016年
3. 雑誌名 英語教育	6. 最初と最後の頁 64-65
掲載論文のDOI（デジタルオブジェクト識別子） なし	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 -

〔学会発表〕 計5件（うち招待講演 5件/うち国際学会 0件）

1. 発表者名 石井雄隆
2. 発表標題 AI・ビッグデータは英語教育をどのように変えるか
3. 学会等名 1210あかりんアワー（招待講演）
4. 発表年 2019年

1. 発表者名 石井雄隆
2. 発表標題 オープンサイエンス時代の外国語教育研究
3. 学会等名 外国語教育メディア学会中部支部第94回支部研究大会（招待講演）
4. 発表年 2019年

1. 発表者名 石井雄隆
2. 発表標題 外国語教育における第二言語ライティング研究・支援ツールの動向
3. 学会等名 作文研究2019（招待講演）
4. 発表年 2019年

1. 発表者名 石井雄隆
2. 発表標題 テクノロジーを活用した英語ライティング研究の近年の動向
3. 学会等名 実践女子大学短期大学部言語文化教育研究センター講演会（招待講演）
4. 発表年 2017年

1. 発表者名 石井雄隆
2. 発表標題 データマイニングを活用した英語ライティング研究
3. 学会等名 大学英語教育学会英語辞書研究会例会（招待講演）
4. 発表年 2016年

〔図書〕 計1件

1. 著者名 石井 雄隆、近藤 悠介	4. 発行年 2020年
2. 出版社 ひつじ書房	5. 総ページ数 164
3. 書名 英語教育における自動採点：現状と課題	

〔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
--	---------------------------	-----------------------	----

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
---------	---------