

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年6月10日現在

機関番号：23503

研究種目：基盤研究（C）

研究期間：2010～2012

課題番号：22520572

研究課題名 タスクに基づくライティングテストの開発とその有用性に関する総合的評価

研究課題名 Development of a task-based writing performance test and comprehensive assessment of the usability

研究代表者 杉田 由仁 (SUGITA YOSHIHITO)

山梨県立大学・看護学部・准教授

研究者番号：70363885

研究成果の概要（和文）：本研究の主要な成果とは、1) 評定の厳しさにおける度合いは異なるが、評定者が一貫した評定を行うことのできる評定尺度および『評定の手引き』を開発することができた、2) 受験者のライティングにおける運用能力を測定する上で、信頼性・妥当性のある評価タスク開発を可能にする「構成概念に基づく言語处理的テスト法」の考え方を確立することができた、3) 評価対象となるライティング能力を効果的に測定することのできる5段階の単特性に基づく評定尺度を開発することができたことであると言える。

研究成果の概要（英文）：The overriding positive contribution of this research is as follows: (a) a clear definition of the writing abilities to be assessed in performance assessment was established, (b) how to develop the task-based writing performance test that has Japanese educational applications was clarified and (c) the writing performance test confirmed usability to some extent through investigating reliability and construct validity. All of them can contribute to the effective writing assessment in the classroom in Japan and to improvements in teaching writing as its washback.

### 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	400,000	120,000	520,000
2011年度	300,000	90,000	390,000
2012年度	500,000	150,000	650,000
総計	1,200,000	360,000	1,560,000

研究分野：英語教育学

科研費の分科・細目：言語学・外国語教育

キーワード：教育評価・測定、ライティング

#### 1. 研究開始当初の背景

(1) 2007年度には、タスクに基づくライティングテスト(Task-based Writing Test, 以後TBWTと呼ぶ)開発の基礎となるライティング評価に関わる先行研究の概観とテスト開発の理論的枠組みおよびライティング能力を構成する概念の検討を行った。この概念規定に基づき、書く領域における言語運用能力を

顕在化させるためのタスクを実際に考案し、これまでに開発された主要な評価尺度(TOEFL, FCE, Michigan)における評価規準を、本研究におけるaccuracyとcommunicabilityの概念規定と照らし合わせながら、それぞれに対応する評価規準を選定して分類・整理、統合することにより新たに6段階の評価尺度を作成した。

(2) 2008 年度には、1) TBWT に関連する要因(評定者、被験者能力、評価タスク)が相互に影響し合う度合いに分析を行う、2) 評価タスクおよび評定尺度の信頼性について検証することを目的として、大学生英語学習者を対象とした予備調査を実施した。高校英語教師 5 名が評定者となり、accuracy, communicability の概念規定に基づいて作成された 6 段階の評価尺度を用いて評定作業を行った。評定結果の分析には FACET ソフトウェア(Linacre, 2006)を活用して、評定者の厳しさの度合い、被験者のライティング能力、タスク難易度を log による同一尺度上に配置するファセットサマリー分析によって行った。その結果から、5 名の評定者の厳しさに関しては同等ではなかったが、全員が一貫した評定を行ったことがわかった。また、各評定者は独自のバイアスを持たずに評定を行うことができ、評定者間の評定が一致した割合(34.6%)は各評定者が“independent experts”として評価を行うことができたことを示す結果となった。しかし、評定尺度に関しては、内的一貫性があり一定の信頼性があることは確認されたが、尺度の等間隔性に課題が残され、6 段階から 5 段階の尺度に改訂する必要があることが示唆された。

(3) 2009 年度には、前年度の予備調査によって明らかにされた評定尺度改訂の必要性について検討を行い、評定尺度の最適化に取り組んだ。具体的にはまず、測定対象となる構成概念(accuracy, communicability)の定義の見直しと評価基準の細分化作業を行い、評定作業のガイドラインである“TBWT Scoring Guide”の改訂版を作成し、本調査実施に向けての本格的な準備を行った。

## 2. 研究の目的

本研究は、特別な訓練を受けたライティング評価の専門家ではなく、実際に教室で指導にあたる中学校・高等学校の英語科教員が、信頼性・妥当性の高いライティングのパフォーマンス評価を行うことができる評価基準・テスト方法の開発手順を確立することを目的とする。

## 3. 研究の方法

### (1) 評定尺度の最適化に関する調査

予備調査の結果に基づき改訂された評定尺度および“TBWT Scoring Guide”等の、ライティングテストとしての信頼性を検証するために、予備調査と同一の評定者に、改訂された評価尺度を活用して、新たなライティング・サンプルの評定を依頼し、評定結果について予備調査と同様のファセット分析を行う。

### (2) 信頼性・妥当性に関する検証

テストとしての信頼性の検証を行うために、新たな評定者に、本調査 1 と同じサンプルを同一の評定尺度を活用して評定作業を行うように依頼する。予備調査を含め 2 回の TBWT に関わった 5 名の高校英語教師と、新規に TBWT に関わった英語教師の評定結果についてファセット分析を行い、信頼性について検証する。また、妥当性の検証に関しては、TBWT の対象者となった大学生学習者に、ETS(Educational Testing Service)によって開発された「オンライン自動採点システム(Criterion)」によるエッセイライティングのテストの受験機会を設定し、客観的な評価指標に基づいて与えられる 6 段階の全体的評価法のスコアとの相関から併存妥当性について検討を行う。

### (3) 実用性に関する検証

TBWT の教育現場における実用性について検証を行うために研究協力校に依頼し、高校生および大学生学習者を対象とした一定規模の受験者によるタスクに基づくライティングテストを実施する。調査においては、評価対象がライティング能力のレベルが異なる高校生および大学生学習者となるので、それぞれのレベルに合わせた評価タスクを考案し、評価基準・尺度、実施形態などについても、これまでに実施した調査結果を参考にして修正・改善を加えて実施する。あわせて対象となる学生には、妥当性検証の基礎データとなる「オンライン自動採点システム(Criterion)」によるエッセイライティングのテストの受験機会を設定する。さらに、TBWT を受験した高校生・大学生学習者を対象

に、テストの実施形態、内容、方法等に関する質問紙調査を依頼し、評定に関わった英語教師に対しても、評定基準および評定尺度の実用性、評定作業に要する時間、労力、負担などに関わる質問項目によって構成される調査票を作成し、回答を依頼する。

#### (4) 評価タスクの考案と追調査

(3)の調査において、本研究の accuracy と communicability という2つの側面からライティングにおける言語運用能力を測定・評価することの信頼性、妥当性、実用性が実証されたところで、応用的研究への準備段階として、それぞれの測定に活用し得る複数の評価タスクを考案して追調査を実施し、その結果についてタスクと尺度を2つの facet として配置する FACETS 分析を行い、それぞれがスコアの分散にどのような影響を与えているかについて検証を行う。

#### (5) コンピューター利用による自動採点システム構築に向けての検討

複数の評価タスクによる追調査においても、TBWT のテストとしての信頼性、妥当性、実用性が実証され、結果の総合的な判断により、本研究が目的とするタスクを測定手段とした「書く」領域における言語運用能力評価およびタスクに基づくライティングテストの有用性が確認された場合には、一定規模の本格的な公開テストの実施に向けて、評価方法としての客観性および有用性をより向上させる取り組みが必要となる。特に評定については、実施規模が拡大し、受験者が増えれば増えるほど迅速な処理方法を用意しておかなければ、テストの実施は困難になることが予想される。そこで、近年著しい進歩を遂げている「コンピューターによる自動採点システム」を活用することにより、評定者要因による影響を除外し、より客観的で実用的なシステム構築に着手する。タスクに基づくライティングテストの検証作業と並行して採点システムの開発を計画する。

### 4. 研究成果

本研究の主要な成果とは、1) 評定の厳しさにおける度合いは異なるが、評定者が一貫

した評定を行うことのできる評定尺度および『評定の手引き』を開発することができた、2) 受験者のライティングにおける運用能力を測定する上で、信頼性・妥当性のある評価タスク開発を可能にする「構成概念に基づく言語处理的テスト法」の考え方を確立することができた、3) 評価対象となるライティング能力を効果的に測定することのできる5段階の単特性に基づく評定尺度を開発することができたことであると言える。

また、これからの日本の学校におけるライティング指導に対しては、1) ライティングによる運用能力評価において、評価対象となる能力を明確に示すことができた、2) タスクに基づくライティング・テストの具体的な開発手順を確立することができた、3) 特別な訓練を受けたライティング評価の専門家ではなく、実際に教室で指導にあたる中学校・高等学校の英語科教員が、タスクを測定手段として行う信頼性・妥当性の高い運用能力評価を行うことができる評価基準・テスト方法を開発することができたことにより、ライティングの評価改善の面で多少なりとも貢献ができるのではないかと考えられる。

さらに、上記の研究成果から導かれる教育的唆を3つの観点から述べる。まず、これからの日本の学校において、ライティングは生徒のコミュニケーション能力養成を目的として指導されるべきであるということである。そのためには、ライティングの指導内容として、「正確さを志向する活動」と「流暢さを志向する活動」の両方を配置する必要がある。その根拠として、Bachman and Palmer (1996) による文法能力とテキスト能力によって構成される「構成能力・知識」と機能的な能力と社会言語学的能力によって構成される「語用論能力・知識」と密接に関わる TBWT の構成概念である「正確さ」と「伝わりやすさ」が、本研究における複数の調査において、一定の信頼性・妥当性を担保して測定されたことを指摘することができる。この測定結果は、取りも直さず、生徒のライティング能力はこれらの言語知識・能力によって構成されていることを裏づけるものであり、コミュニケーション能力養成を目的として行われるライティング指導に際しては、「正確さ」と

「流暢さ」の両面に関わる指導内容が求められるということを示唆する。そして、TBWT の評価基準に明示された「言語能力特性」に基づくならば、文章構成力や言語的正確さの指導および伝達内容の質や情報伝達の効果を高める指導を行うことがライティングにおけるコミュニケーション能力を養成する有効な手立てとなると考えられる。

次に、これからの日本の学校においては、より積極的にライティングの運用能力評価を取り入れるべきであるということである。これまでの日本の学校におけるライティング指導は、従来の「形式重視」の指導から「内容重視」の指導へと変容しつつあるが、その評価法は文法や語法に関する多肢選択問題や文の結合・誤文訂正問題などの間接テストが主流であり、これといった進展が見られていない。その理由として、直接テストによるパフォーマンス評価は採点に時間がかかることや評定作業の負担が大きいことなどがあげられる。しかし、本研究において開発された TBWT によれば、テスト実施時間は概ね 30 分程度であり、標準の授業時間内で十分に実施可能である。分割して実施すれば、10 分あるいは 20 分で小テスト的に行うことも可能である。また、評定については単特性の 5 段階尺度による全体的評価方式であるため、多忙な現場教師にも十分に利用可能であると考えられる。つまり TBWT の内容構成は、生徒に直接ライティングを行わせ、実際のパフォーマンスにより運用能力を測定しようという教師の意思により、十分に実用化が可能であり、教室へのライティング運用能力評価導入を実現するものであると言える。

第 3 の教育的示唆として、英語教師教育プログラムの一環として、評定者トレーニングを導入すべきである。本研究においては、5 名の高校教師と 15 名の中学校教師に研究協力者として評定を依頼した。校種や実践経験等を考慮して評定者グループを組織して評定作業を行っていただいた結果、いずれの評定セッションにおいても評定者の信頼性や一貫性が確認された。しかし、評定者間の厳しさの違いや各評定者固有のバイアス傾向が認められ、評定者トレーニングの必要性が示唆された。つまり、英語教師としての実践

経験はライティングの評定に影響を与える一要素に成り得るとは言えるが、経験を積み重ねるのみでは主観的判断に左右される評定の域を出ない (Schaefer, 2008) ということである。Richard and Farrell (2005) が、「トレーニングを行うことは指導に必要不可欠な基本的概念や原理の理解および教室における実践力の向上につながる (p. 3)」と述べている通り、評定者トレーニングを実施することにより、TBWT の基本原理を理解し評定方法についてより習熟することが可能になると考えられる。さらに、教師教育プログラムの一環として実施されるならば、TBWT における評定一貫性の向上が図られるという効果に留まらず、ライティングのパフォーマンス評価そのものに対する理解が深められ、教室におけるライティングの指導改善にも波及するものと考えられる。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 5 件)

- ① Yoshihito Sugita, Comparability of accuracy and communicability tasks: Are they all equally difficult? *Proceedings of the 17<sup>th</sup> Conference of Pan-Pacific Association of Applied Linguistics*, 査読有、2012、81-82
- ② Yoshihito Sugita, The effects of rater training on raters' severity, consistency, and biased interactions in a task-based writing assessment, *日本語テスト学会誌*、査読有、第 15 号、2012、61-80
- ③ 杉田由仁、ライティング評価における評定者の行動分析と評価基準の妥当性検証、*JACET 関東支部学会誌*、査読有、2012、14-26
- ④ Yoshihito Sugita, Differences in raters' severity, consistency, and biased interactions between trained and untrained raters, *Proceedings of the 16<sup>th</sup> Conference of Pan-Pacific Association of Applied Linguistics*, 査読有、2011、73-78
- ⑤ Yoshihito Sugita, Reliability and validity of a task-based writing performance assessment for Japanese learners of English, *日本語テスト学会誌*、査読有、第 13 号、2010、21-40

[学会発表] (計4件)

- ① Yoshihito Sugita, Comparability of accuracy and communicability tasks: Are they all equally difficult? The 17<sup>th</sup> Conference of PAAL, 2012年8月22日(北京・中国)
- ② Yoshihito Sugita, Differences in raters' severity, consistence, and biased interactions between trained and untrained raters, The 16<sup>th</sup> Conference of PAAL, 2011年8月9日(香港・中国)
- ③ 杉田由仁、ライティング評価における評定者の行動分析と評価基準の妥当性検証、第5回大学英語教育学会・関東支部大会 2011年6月26日(大東文化大学)
- ④ Yoshihito Sugita, Reliability and validity of a task-based writing performance assessment of L2 writing, The 44<sup>th</sup> Annual Conference of IATEFL, 2010年4月11日(ハロゲイト, 英国)

[図書] (計1件)

- ① 杉田由仁、大学教育出版、日本人学習者のためのタスクによるライティング評価法—構成概念に基づく言語处理的テスト法— 2013、232

[その他]

ホームページ等

<http://homepage3.nifty.com/~sugita/index.html>

## 6. 研究組織

### (1) 研究代表者

杉田 由仁 (SUGITA YOSHIHITO)  
山梨県立大学・看護学部・准教授  
研究者番号：70363885

### (2) 研究分担者 なし

### (3) 連携研究者 なし