

令和 4 年 6 月 16 日現在

機関番号：33501

研究種目：基盤研究(C) (一般)

研究期間：2017～2021

課題番号：17K03030

研究課題名(和文) スローラーナー向けの英作文分析的評価基準の作成

研究課題名(英文) Development of Analytic Evaluation Criteria for Novice EFL Writers

研究代表者

馬場 千秋 (Baba, Chiaki)

帝京科学大学・教育人間科学部・教授

研究者番号：50465374

交付決定額(研究期間全体)：(直接経費) 2,700,000円

研究成果の概要(和文)：本研究では、英語を不得意とするスローラーナー向けの英作文分析的評価基準の開発を目的とし、既存の評価基準の検討、適した評価項目、総語数の扱いについて検証を行った。その結果、スローラーナーの英作文を評価するためには、内容、語彙、文法、総語数の4つの項目が適していることが明らかとなった。また、限られた時間内に英作文を書くことが容易であるのは、TOEIC Bridgeスコア130以上の学習者であることがわかった。これは、文法上のエラーが減り、「独立した書き手」となるレベルとTOEFL Writing Scoring Guideを使用できるレベルと一致していることを検証することができた。

研究成果の学術的意義や社会的意義

本研究では、英語を不得意とするスローラーナー向けの英作文分析的評価基準を開発した。既存の評価基準を用いた場合、学習者の書いた英作文の評価が低いところに偏るため、どのような点を指導により強化していくかが見えにくかった。今回開発した評価基準は、スローラーナーが書いた英作文を評価するために適した評価項目を備え、小さな進歩も評価することができる。今後、授業の中で英語を書く活動を積極的に取り入れたときに、この評価基準を使うことで、学習者にできている事柄と強化すべき事柄を伝えると同時に、自己表現をしたことによる達成感を感じさせることが可能となる。

研究成果の概要(英文)：This study develops new analytic criteria for novice EFL writers. To complete the task, I analyzed whether the ESL Composition Profile (Jacobs, et al., 1981) can be used for novice writers and found that the profile was not suitable for novice learners. Next, I investigated the appropriate evaluation items. The results showed that four items, that is, content, vocabulary, grammar, and word count are the best items for novice EFL writers. I also found that writers with TOEIC Bridge scores over 130 could write English compositions fluently and accurately, and their compositions can be evaluated with the TOEFL Writing Scoring Guide.

研究分野：英語教育学

キーワード：英作文 分析的評価基準 スローラーナー

## 1. 研究開始当初の背景

英語で書かれた英作文を評価する場合、全体的評価、分析的評価、primary trait scales の3つが挙げられる。この中で primary trait scales については、特定のライティングタスクにのみ適しているため、一般的には全体的評価と分析的評価のいずれかを使用することが適している(Weigle, 2002)。全体的評価には、TOEFL Writing Scoring Guide や Cambridge English Language Assessment の各テストや IELTS で用いられているスケールが挙げられる。一方、分析的評価には Jacobs, et al. (1981)の ESL Composition Profile、Weir (1988)の Test in English for Educational Purposes (TEEP)などがある。しかし、これらの評価基準の多くは、留学を目指す学習者やある程度の英語力を持った学習者のライティングパフォーマンスを測るために用いるには適しているが、英語を不得意とするスローラーナーのライティングパフォーマンスを測ろうとすると、低い評価が与えられ、個々の学習者のライティング力を細分化して評価することができない。そのことを踏まえ、Baba(2016)では、授業の中で書かせた英作文を教師が短時間で採点できるようにするため、全体的評価に着目し、スローラーナー向けの評価基準を作成し、高い信頼性と妥当性を得ることができた。この研究の中で作成した全体的評価基準の信頼性と妥当性の高さを検証するために、分析的評価基準では定評のある Jacobs, et al. (1981)の ESL Composition Profile との比較を行った。その結果、ESL Composition Profile はスローラーナーの英作文を評価することはできるが、content, organization, vocabulary, language use, mechanics の5つの項目の評価点が均一ではなく、評価そのものに偏りが見られることが問題点として明らかとなった。分析的評価については、Hamp-Lyons(1991)も、複数の側面から分析的に評価する場合、それぞれの側面は均一に評価されるべきであることを指摘している。

日本人英語学習者を対象にした英作文評価基準の作成を試みた研究の1つとして、Jacobs, et al. (1981)の ESL Composition Profile を参考にして作成された Nakanishi(2006)がある。この評価基準は content, organization, vocabulary, language use, mechanics の5つの観点から成り立っており、5段階で評価を行うものである。しかし、各観点についての説明が簡単すぎるため、採点者にとって使いやすい採点基準にするには観点についてより詳しい説明が必要である。

ループリックで評価をする試みとしては、Nishijima, et al. (2007)が、パイロット版として、content and idea development, organization, grammar, vocabulary, mechanics の5つの観点の基準を作成した。これを改良したものが Kinshi, et al. (2011)の Rubric 2008 であり、一般化可能性理論を用いて信頼性を検証した。比較的高い信頼性が得られたが、学習者の英語力が多様であるにもかかわらず、採点者の評価は同じような点数であったため、再改良を行い、Rubric2009 を作成した。content and idea development, organization, grammar, vocabulary, mechanics の5つの観点から成り、4段階で評定を行うものである。このループリックを用いることで、ターゲットを絞った特定の観点からの評価ができるが、点数配分などはまだまだ考慮すべきものである。大久保 (2006)も、ESL Composition Profile をもとにして、content, organization, vocabulary, language use の4つの観点から成る分析的評価基準を作成したが、信頼性が低く、更なる修正が必要である。

上記の先行研究から、英語を不得意とするスローラーナーに英作文指導を行う場合、英作文課題を課す回数が多ければ、Baba(2016)の全体的評価が適しているが、診断的評価を行い、弱点強化と英語を書くことによるコミュニケーション能力を伸ばすことを目的とした場合、分析的評価を併用することがより効果的である。そこで、本研究では、スローラーナー向けの分析的評価基準を作成することを試みた。なお、この基準を作成するにあたり、英語を不得意とする大学生だけでなく、英語初級学習者である中学生や高校生にも使用することが可能なものとすることを目指した。

## 2. 研究の目的

本研究の目的は、英語を不得意とするスローラーナー向けの英作文分析的評価基準を作成することである。そのために、

- (1) 既存のライティング分析的評価基準はスローラーナーに使用できるか。
- (2) スローラーナー向けの分析的評価基準に必要な観点は何か。
- (3) パイロット版分析的評価基準の信頼性検証
- (4) 英文産出量は分析的評価基準として妥当か。
- (5) 完成版分析的評価基準の信頼性検証

を行うこととする。

## 3. 研究の方法

本研究は、5つの研究からなる。

### 3. 1 既存のライティング分析的評価基準はスローラーナーに使用できるか

### 3.1.1 被験者

日本人英語教員 9名

### 3.1.2 研究手順

研究代表者担当クラスで収集した大学生英語学習者の英作文データのうち、TOEIC Bridge スコア 72-146 の学習者のデータを 50 編無作為抽出し、9 名の日本人英語教員に ESL Composition Profile (Jacobs, et al, 1981)を用いて採点をしてもらった。その後、使用した Profile の使いやすさや採点項目の重要度についてのアンケートに回答してもらった。

採点結果については、採点者間信頼性が高いことを確認した上で、9 名の採点結果の平均値と TOEIC Bridge スコアの相関を算出した。さらに、ESL Composition Profile で採点可能な学習者のレベルおよび必要な採点者数、必要な評価基準について、検証した。

### 3.2.3 研究結果

9 名の採点者間信頼性は高く、クロンバック 係数も.922 で、効果量も中であった。分析的評価の項目それぞれにおいても、採点者間信頼性は高かった。ESL Composition Profile の採点結果と TOEIC Bridge スコアの相関はクロンバック 係数では.571、効果量は中であり、それぞれの評価項目を見ても、相関はさほど高くはないという結果であった。Baba(2016)では、評価基準を使うことのできる学習者の英語力を見出すことができたが、本研究では閾値を見出すことができなかった。重回帰分析を行ったところ、Content, Organization, Vocabulary, Language use, Mechanics の 5 つの指標のうち、Mechanics のみが予測可能という結果となった。

一般化可能性理論を用いて、評価基準数と評価者数を算出したところ、評価基準が多ければ多いほど、より多くの評価者がいなければ、信頼性が高くはならないこと、また、一人の評価者が採点する場合もぶれが生じる可能性が示唆された。

本研究で使用した評価基準は、スローラーナーを評価するには使い勝手がよくないこともアンケート結果から明らかとなった。また、採点項目の重要度については、内容、語数、文法を重視する評価者が多く、文章校正については、スローラーナーの英作文を評価する場合には、あまり重視されていないこと、そして評価するのも難しいことが明らかとなった。

## 3.2 スローラーナー向けの分析的評価基準に必要な観点は何か

### 3.2.1 被験者

日本人英語教員 7名

### 3.2.2 研究手順

研究代表者担当クラスで収集した大学生英語学習者の英作文データのうち、TOEIC Bridge スコア 72-146 の学習者のデータを 30 編無作為抽出し、7 名の日本人英語教員に 8 つの観点（内容、文章構成、語彙、語数、文法、パンクチュエーション、日本語の翻訳によるエラー、課題の完成度）を 6 段階で評価してもらい、評価後、8 つの観点が評価しやすいかどうかのアンケートに回答してもらった。7 名の採点者間信頼性を検証後、評価点の平均値を出し、それぞれの項目の評価と英作文を書いた大学生の TOEIC Bridge のスコアとの関係性を検証した。さらに、必要な評価項目について、検討を行った。

### 3.3.3 研究結果

7 名の採点者間信頼性は、語数、語彙、文法、内容と完成度においては高く、クロンバック 係数も高い値が算出された。その一方で、文章構成、パンクチュエーション、翻訳によるエラーの信頼性は低い結果となった。また、8 つの項目の相関は高い結果となった。

一般化可能性理論を用いて、評価基準数と評価者数を算出したところ、5 つの項目を設定した場合は 3 名以上の評価者、3 項目の場合は 4 名以上の評価者がいれば、高い信頼性を得られることが明らかとなった。この結果、語数、語彙、内容、文法がスローラーナー向けの評価基準には適した評価項目であることが立証された。

## 3.3 パイロット版分析的評価基準の信頼性妥当性検証

### 3.3.1 被験者

日本人英語教員 7名

### 3.3.2 研究手順

研究代表者担当クラスで収集した大学生英語学習者の英作文データのうち、TOEIC Bridge スコア 84-168 の学習者のデータを 40 編無作為抽出し、7 名の日本人英語教員に 5 つの観点（内容、語彙、文法、語数（定義されているもの）、実語数）からなるパイロット版の分析的評価基準で評価してもらった。評価後、5 つの観点が評価しやすいかどうかのアンケートに回答してもらった。7 名の採点者間信頼性を検証後、評価点の平均値を出し、それぞれの項目の評価と英作文を書いた大学生の TOEIC Bridge のスコアとの関係性を検証した。さらに、パイロット版の評価基準で改良が必要な項目は何か、改良すべき点は何かについて、アンケート結果をもとに、インタビューを行い、パイロット版の評価項目のうち、評価の定義の説明が曖昧な項目について、評価のしにくさの要因および、説明に関する記述をどのように修正していけばよいのか、検討を行った。

### 3.3.3 研究結果

パイロット版を用いた 7 名の採点者間信頼性は各項目の点数を合計した場合、信頼性が高

くなり、効果量も大であった。個々の項目を見た場合、内容、語数は信頼性が比較的高めであったが、語彙、文法は評価のしにくさが生じ、採点者間信頼性も中程度となった。そのため、アンケートにおいて、パイロット版を用いて評価をしやすいかどうかを問うたところ、実語数以外は 6 段階で 5 以下の平均値であった。特に語彙と文法については、定義の説明が曖昧なため、2 つの点数のうち、どちらにしたらよいかわからないなど、問題点も散見された。

また、一般化可能性理論を用いて、評価基準数と評価者数を算出したところ 3 つの項目で評価する場合でも、4 名の評価者でなければ、高い信頼性が得られないという結果であった。

上記の結果より、パイロット版については、定義の説明で曖昧な記述を明確な記述に修正し、改良版を作る必要があることが明らかとなった。そこで、評価基準が使いやすいかを問うたアンケートの自由記述欄にコメントを記載してくれた 5 名の記述を検討し、よりよい評価基準にするにはどのような説明にすべきか、自由記述をもとにインタビューを行った。その中で指摘が多かった記述としては、似たような記述が複数個所（点数）に見受けられ、どちらの点数にすべきかがわかりにくいことや、1 つの英作文の中にみられるいくつかの現象が、評価の点数の記述を見ると、バラバラに入り込んでいるため、評価できないこと、「適切」「散見される」「少し不足」など、評価者によって解釈が異なる記述があり、評価をするときにも記述の解釈で迷ってしまうことなどが挙げられた。語数については、実語数と定義により判断するものと双方の基準を設けたが、実語数のほうが確実であり、評価がしやすいとのコメントを得た。このインタビュー結果をもとに、評価基準を改良し、最終版を作成していくための検討を行うこととした。

### 3.4 英文産出量は分析的評価基準として妥当か

#### 3.4.1 被験者

東京都の私立大学教育人間科学部学校教育学科 1 年生 125 名  
(TOEIC Bridge スコア平均: 112)

教員養成課程に在籍し、小学校、中高理科、中高保健体育、中高英語のいずれかの教員免許状を取得することを目指している学生である。

#### 3.4.2 研究手順

被験者は英語力を確認するため、TOEIC Bridge の模擬試験を受験した。その後に受講している授業の 20 分間で Timed Writing を実施した。用紙と 4 色ボールペン（黒、赤、青、緑）を配布し、5 分経過するごとに指定した色に変え、英作文を書かせた。課題は旅行に行きたい場所について、理由、行って何をしたいのか、SNS に旅行記を載せることを想定するなど、いくつか条件を提示した上で、できるだけ詳しく書くように指示をした。書いていて間違えた場合は、線を引いて間違えた箇所を消して、次の文を書くようにさせた。

被験者が書いた英作文は、色も英作文が書かれた色に合わせて Microsoft Word に打ち込み、データとし、総合数と 5 分ごとの語数を記録した。5 分ごとのそうご数の伸びについては、繰り返しのある一元配置の分散分析を行い、統計的に有意かどうかを検証した。また、20 分間のどの段階での語数が多いか、分析を行った。

英作文そのものは、日本人英語教員 2 名が語彙、文法、内容の 3 つの観点の分析的評価 (Jacobs, et al, 1981 の抜粋、改良版を使用) および Writing Criteria for Novice Learners (Baba, 2016) を用いた全体的評価を行った。

英語力と英文産出量の関係を見るために、総語数と TOEIC Bridge スコアの相関を取り、英作文の評価と総語数の関係を見るために、分析的評価、全体的評価と総語数の相関を取り、分析を行った。

#### 3.4.3 研究結果

##### 3.4.3.1 Timed Writing

5 分ごとの Timed Writing の結果であるが、最初の 5 分間に書いている語数が多く、その後、英文産出量は減っていく傾向にある。レベル別に見た場合、TOEIC Bridge スコア 110 以下の学習者に顕著である。英語が不得意なスローラーナーの場合、「何をどのように書いたらいいのかわからない」と感じる学習者が多く、日本語でも作文を書いた経験が少ないこともあり、時間が経つにつれ、完全に手が止まってしまい、最後の 5 分はほとんど書いていない、という学習者も見受けられた。一方、TOEIC Bridge スコア 130 以上の学習者の場合、はじめの 5 分の総語数が多く、その後、英文産出量は少し減ってくるが、英文を書き続け、校正する傾向にあるため、時間を追うごとに英文産出量が極端に減るということはない。馬場 (2011) において、研究代表者はスローラーナーのエラーの傾向を分析し、「独立した書き手」となるレベルを TOEIC Bridge スコア 130 以上ということを指摘し、また、Baba (2016) でも TOEFL Writing Scoring Guide が使えるレベルも TOEIC Bridge スコア 130 であることを明らかとしたが、英文産出量においても、TOEIC Bridge スコア 130 が指標となることが立証された。

総語数と英語力の相関は、 $r=.623$  と比較的高い相関が見られる。英語力と英文産出量が比例することも検証された。

総語数の伸びについては、全体の 5 分毎について、繰り返しのある一元配置の分散分析を行い、20 分間の変化を見たところ、 $F(2.692, 333.860) = 38.029$ ,  $p < .001^{**}$  であり、統計的にも有意な差があることが明らかとなった。

### 3.4.3.2 評価

英作文は2名の英語教員が採点したが、採点者間信頼性については、全体的評価が  $N = 125$   $r = .916$   $**p < .001$ 、クロンバック 係数が.955、効果量が大であった。内容については、 $N = 125$   $r = .887$   $**p < .001$ 、クロンバック 係数が.925、効果量が大、語彙は  $N = 125$   $r = .797$   $**p < .001$ 、クロンバック 係数が.886、効果量が大、文法は  $N = 125$   $r = .756$   $**p < .001$ 、クロンバック 係数が.860、効果量が大であった。したがって、採点者間信頼性は高いと言える。

総語数と英作文評価は、全体的評価、内容、語彙、文法、全てにおいて相関が高い。レベル別にみても、TOEIC Bridge スコア 120 以上の総語数が多く、130 以上になるとさらに総語数の伸びは大きくなっていることがわかる。英作文評価についても、TOEIC Bridge スコア 120 以上とそれ以下ではかなりの差が見られる。また、TOEIC Bridge スコア 130 以上の学習者については、全体的評価と内容において、差が見られる。英語力が高くなればなるほど英作文の完成度も高くなる傾向にある。その一方で、語彙と文法については、大きな差は見られない。初級から中級のレベルの学習者の場合、発表語彙があまり多くないことや複雑な文法項目を使いこなせないことが要因であると考えられる。

本研究の結果より、総語数（英文産出量）は、スローラーナー向けの分析的評価基準として適切であると考えられる。全体的評価基準の場合は、他の基準と合わせて考えていく必要があるため、項目に含めることが難しいが、分析的評価基準の場合は、単独で評価できる項目として設けておくことが望ましいと言える。

### 3.5 完成版分析的評価基準の信頼性妥当性検証

#### 3.5.1 被験者

日本人英語教員 5名

#### 3.5.2 研究手順

研究代表者担当クラスで収集した大学生英語学習者の英作文データのうち、TOEIC Bridge スコア 44-156 の学習者のデータを 60 編無作為抽出し、5名の日本人英語教員にパイロット版から改良をした、4つの観点（内容、語彙、文法、総語数）の分析的評価基準と Baba (2016) の全体的評価基準で評価をしてもらった。5名の採点者間信頼性を確認後、5名の評価の平均値、作文を書いた学習者の TOEIC Bridge、総語数の相関を取った。

#### 3.5.3 研究結果

4つの分析的評価項目、全体的評価、TOEIC Bridge および総語数の相関はいずれも高い傾向にあり、クロンバック 係数は.512、効果量は大であった。

一般化可能性理論を用いて、採点者数を算出したところ、2項目以上の採点基準の場合、2名以上の採点者がいれば、信頼性があることが明らかとなった。

本研究において作成した分析的評価基準によって、高い信頼性を得ることができることが立証された。

### 4. 研究成果

本研究では、英語を不得意とするスローラーナー向けの英作文分析的評価基準の開発を目的として、既存の評価基準の検討からはじめ、適した評価項目、総語数の扱いについても検証を行った。その結果、スローラーナーの英作文を評価していくためには、内容、語彙、文法、総語数という4つの項目が適していることが明らかになることができた。また、評価基準作成の過程の中で、限られた時間内に英作文を書くことが容易であるのは、TOEIC Bridge スコア 130 以上の学習者であることを見出すことができた。これは、文法上のエラーが減り、「独立した書き手」になりつつあるレベルと TOEFL Writing Scoring Guide を使用できるレベルと一致しており、英語を書くことが苦にならなくなるのがこのレベルだということが複数の研究結果を通じて検証することができた。

今後の課題としては、TOEIC Bridge スコア 130 未満の学習者と 130 以上の学習者、それぞれの英語ライティング力を伸ばすための指導法や教材について検討していくことである。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計7件（うち招待講演 0件 / うち国際学会 5件）

1. 発表者名 Baba, Chiaki
2. 発表標題 Reconsidering Analytic Evaluation Criteria for Novice Japanese EFL Learners
3. 学会等名 The 17th Asia TEFL International Conference and The 6th FLLT International Conference (国際学会)
4. 発表年 2019年

1. 発表者名 馬場 千秋
2. 発表標題 スローラーナーのTimed Writingにおける英文産出量と分析的評価
3. 学会等名 第45回全国英語教育学会弘前研究大会
4. 発表年 2019年

1. 発表者名 Baba, Chiaki
2. 発表標題 Developing a New Analytic Evaluation Criteria for Novice EFL Learners: A Pilot Study
3. 学会等名 Symposium on Second Language Writing 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Baba, Chiaki
2. 発表標題 How Should Practitioners Create Rapport with Their Students Through Writing Feedback?
3. 学会等名 The 1st JACET Summer (45th) and English Education (6th) Joint Seminar (Kyoto, 2018)
4. 発表年 2018年

1. 発表者名 Baba, Chiaki
2. 発表標題 Reconsidering How to Evaluate Novice EFL Learners' Composition Analytically
3. 学会等名 2017 ALAK International Conference (国際学会)
4. 発表年 2017年

1. 発表者名 Baba, Chiaki
2. 発表標題 How Many Criteria are Appropriate for Analytically Evaluating English Composition?
3. 学会等名 26th ETA International Symposium and Book Fair on English Teaching (国際学会)
4. 発表年 2017年

1. 発表者名 Baba Chiaki
2. 発表標題 How to Evaluate Novice EFL Students' English Compositions: Development of New Analytic Evaluation Criteria
3. 学会等名 The 19th Asia TEFL International Conference 2021 (国際学会)
4. 発表年 2021年

〔図書〕 計0件

〔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
---------	---------