

英語運用力が高い大学生による
E-learning エッセイライティング活動
-*Criterion Writing Evaluation* の実践紹介と効果検証-
Essay Writing Activities on E-learning
by University Students with High English Proficiency
-Introducing *Criterion Writing Evaluation* and Investigating Its Efficacy-

羽山恵*

Megumi Hayama

Email: megkimura@dokkyo.ac.jp

本稿は、外国語学部において英語を専攻する大学生、とりわけ英語力が比較的高い学生たちに対して行った、E-learning による英語エッセイライティング活動実践の紹介である。使用している E-learning 教材は、Educational Testing Service (ETS) が開発した *Criterion* で、自動的に短時間で英語エッセイの総合的評価および診断的フィードバックを出力し、診断的フィードバックの項目が多岐にわたっていることが魅力である。また、継続的なその活動の有用性を、第 2 言語習得研究分野で提唱されているライティング能力の評価指標を用いて検証した。その分析方法には、コーパス言語学で開発・提案された言語情報処理技術を応用した。分析の結果、向上が見られるのはライティングの「流暢さ」および使用する「語彙の洗練度」だということがわかった。

The present paper introduces English essay writing activities on E-learning by English major university students with relatively high proficiency. The E-learning program is *Criterion Writing Evaluation*, which has been developed by the research group of Educational Testing Service (ETS). It automatically and quickly provides holistic scores and diagnostic feedbacks with various aspects. The paper also investigates whether the repeated activities contribute to the development of the learners' second language (L2) writing performance. The indices are from L2 writing studies and the methodology is applied from corpus linguistics studies. The obtained results indicate that 'fluency' and 'the use of sophisticated words' reflect the development.

*: 獨協大学外国語学部

1. はじめに

本論は、外国語学部において英語を専攻する大学生、とりわけ英語力が比較的高い学生たちに対し、継続的なE-learningによる英語エッセイライティング活動を課し、その有用性を検証したものである。

ここで用いられるE-learning教材は、米国のNPOであるEducational Testing Service (ETS)が開発した自動英作文評価システムを搭載したCriterion¹である。獨協大学外国語学部英語学科では、2006年より1年生全員に『E-learning』の受講を課しており、入学時のプレースメントテストで上位クラスに配属された学生たち(Aグループ)が、毎年Criterionによる英語エッセイライティング活動を行っている²。Criterionによるこの活動を導入した目的は、(1) E-learningによって受講生の自律的英語学習を推進するため、(2) 英語エッセイを大量に書く機会を受講生に与えるため、(3) 大人数に何度も英語エッセイを書かせる活動においても、指導者(担当教員)の添削のための労力が省かれるためである。

Aグループに属する学生はおおよそ50名で、大学入学時のTOEIC得点が750点以上の者から成り、その中には帰国子女も多数含まれる。概して帰国生たちは、高い英語スピーキング能力を有しているが、談話構成に留意したアカデミックライティングを不得手にしている傾向がある。また、日本国内の英語学習・教育によって高い能力を有した学習者は、語彙力や文法正確性、リーディングなどに優れているが、アウトプット能力、とりわけ大量かつ流暢に発話あるいは作文する力が不足している傾向がみられる。このような学習者たちから成る本論における学習者グループが、継続的にE-learningによるエッセイライティング活動を行い、英語能力のどのような観点に伸びが見られるかを検証するのが、本論の目的である。

また、受講生が書いた英語エッセイはコンピュータ可読形式で収集・保存されるため、それはいわゆる「学習者コーパス(learner corpus)」として貴重な資料になり得る。学習者コーパスを分析することによって、当該の学習者の英語運用力を測ることが可能となり、さらにその分析に言語情報処理技術を応用すれば、短時間に労少なく目的を達成することも可能となる。本論では、収集したエッセイデータを学習者コーパスとみなし、コーパス言語学分野で

これまで開発、使用されてきた言語情報処理技術の応用も、あわせて試みることにする。

2. Criterionによるエッセイライティング活動

2.1 Criterionの概要

Criterionは上述の通り、米国ETSが開発したe-rater³という自動添削システムを搭載した、英語エッセイライティング活動のためのE-learning教材である(Monaghan & Bridgeman, 2005⁴)。日本国内への販売においては、国際教育交換協議会(CIEE)日本代表部が代理販売を行っている⁴。使用にあたっては、年間のライセンス契約を結び、指定されたウェブサイトにてログインして活動を開始する。インターネット上のサービスのため、専用のサーバーやソフトウェア等のインストールは不要である。

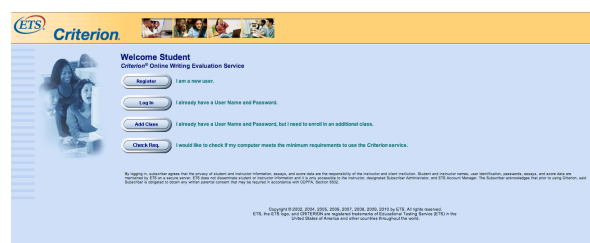


図1 Criterionログイン画面(学習者用)⁵

CriterionにはAdministrator(管理者)、Instructor(教師)、Student(学習者)の3つの役割が設けられており、AdministratorおよびInstructorは自分の学習者を対象としたクラスを作り、彼らに最適なトピックを選定し、課題の開始日・時間や課題の提出締め切り日・時間を設定することができる。その際、一つの課題に対して学習者が何度再提出することができるかの設定も可能である⁶。

学習者は、与えられたトピックに従って英語エッセイを画面(またはMS Word等ワープロソフト)にタイプし、完成したものを提出(submit)する。その後、およそ20秒で、(1)エッセイの6段階による総合的評価(holistic evaluation)と(2)文法(grammar)、用法(usage)、書式(mechanics)、形式(style)、構成(organization and development)についての診断的フィードバック(diagnostic feedback)が表示される。総合的評価の「6」は最も優れており「Excellent」を意味する。以下それぞれ

¹ <http://www.ets.org/criterion/>

² 2006年度当初はAグループのみにCriterionによるエッセイライティング活動が課せられていたが、2008年度から2011年度現在は、レベル別グループに関係なく、全ての英語学科生が活動を行っている。

³ <http://www.ets.org/erater/about/>

⁴ <http://www.cieej.or.jp/toefl/criterion/index.html>

⁵ <https://criterion28.ets.org/cwe/student/>

⁶ Criterionの年間契約料は、期間中の最大提出回数によって異なる。

れ、「5」は「Skillful」、「4」は「Sufficient」、「3」は「Uneven」、「2」は「Insufficient」、「1」は「Unsatisfactory」である。一方の診断的フィードバックは、それぞれのカテゴリーにさらに細分化された項目があり、それぞれに関わるエラーの数や分析的数値が表示される。代表的なものを挙げると、「文法」の場合は「主語と動詞の一致」(Subject-Verb Agreement)や「代名詞の誤り」(Pronoun Error)、「用法」の場合は「冠詞の省略または過剰使用」(Missing or Extra Articles)や「語彙形態の誤り」(Wrong Form of Word)、「書式」の場合は「綴り」(Spelling)や「句読点の付け忘れ」(Missing Final Punctuation)、「形式」の場合は「同一単語の繰り返し」(Repetition of Words)や「短すぎるセンテンスの使用」(Too Many Short Sentences)、そして「総語数」や「センテンス数」、「センテンスの平均語数」などの情報、「構成」の場合は *Criterion* が認識したエッセイの中の「導入部分」(Introductory Material)、「主張」(Thesis Statement)、「主題」(Main Ideas)、「サポーターセンテンス」(Supporting Sentences)、「結論」(Conclusion) が色づけされ指摘される。下記図 2 はフィードバック画面の一例、図 3 はエッセイの構成が色づけされた例である。

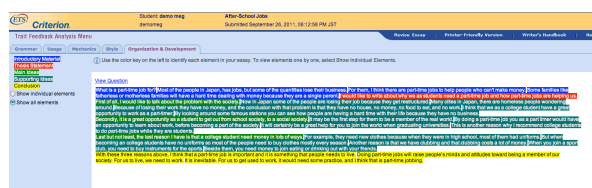


図 2 フィードバック画面の一例

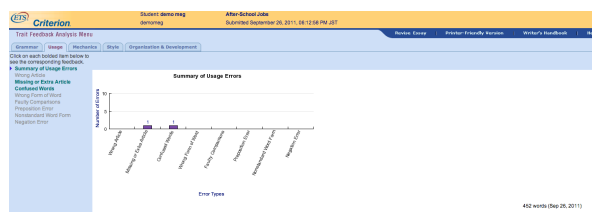


図 3 エッセイの構成に関するフィードバックの一例

2.2 授業におけるエッセイライティング活動

ここで紹介するのは、英語運用能力が高い英語学科 A グループに対する *Criterion* を用いたライティング活動を、最も特徴的に行っていた 2008 年度の授業手順と得られたデータである。一連の活動は以下のように行っていた。

- (1) 対面授業において、担当教員がトピックの提示と課題の指示

- (2) 受講生は 5 日間かけて *Criterion* を用いたエッセイライティング活動に取り組む
- (3) 活動にあたっては、総合的評価と診断的フィードバックを参照しながら、同じものを 3 回まで提出できる (2 回の書き直しと再提出が可能)
- (4) スコア 6 または 5 を取ることを目指す
- (5) 1 週間後の対面授業において、担当教員がクラス全体のスコアの内訳と共通して多かった誤りについての解説を行う (30 分程度)
- (6) 次回のトピックの提示と課題の指示

受講生は通年で 24 回のライティング活動を行った。エッセイのトピックは、予め *Criterion* に収録されているもののうち、GRADE 11 (高校生程度) ~ College Level Second Year (大学 2 年生程度) に分類されているものを選定した。たとえば、After-School Jobs (学生がアルバイトをすることについてどう考えるか)、Experience or Books (書籍から学ぶことと、経験から学ぶことはどのように異なるか) といった論述的 (persuasive) なもの、A Pet's View (動物の視点で自分の家を紹介する) といった記述的 (descriptive) なもの、The Quest (冒険物語) といった物語 (narrative) など、モードにもバリエーションがあった。

以下表 1 は、2008 年度の実施の概要である。

表 1 ライティング活動の時期と参加人数

| 回 | 実施時期 | 人数 | 回 | 実施時期 | 人数 |
|----|------------|----|----|-------------|----|
| 1 | 2008 年 4 月 | 45 | 13 | 2008 年 7 月 | 10 |
| 2 | 2008 年 4 月 | 42 | 14 | 2008 年 9 月 | 41 |
| 3 | 2008 年 4 月 | 46 | 15 | 2008 年 10 月 | 43 |
| 4 | 2008 年 5 月 | 43 | 16 | 2008 年 10 月 | 42 |
| 5 | 2008 年 5 月 | 43 | 17 | 2008 年 10 月 | 37 |
| 6 | 2008 年 5 月 | 44 | 18 | 2008 年 10 月 | 38 |
| 7 | 2008 年 5 月 | 44 | 19 | 2008 年 10 月 | 32 |
| 8 | 2008 年 5 月 | 46 | 20 | 2008 年 11 月 | 37 |
| 9 | 2008 年 6 月 | 45 | 21 | 2008 年 11 月 | 41 |
| 10 | 2008 年 6 月 | 41 | 22 | 2008 年 11 月 | 36 |
| 11 | 2008 年 6 月 | 43 | 23 | 2008 年 12 月 | 30 |
| 12 | 2008 年 6 月 | 39 | 24 | 2008 年 12 月 | 33 |

2.3 収集データ

2008 年度の E-learning 授業において受講生が書いた英語エッセイは、電子データとして収集・保存された。その総数は 951 エッセイ、総語数にして 379,215 語となった。ひとつのトピックあたりの平

均総語数は398.7語(最小値:283語、最大値:474語)であった。また、トピックあたりのCriterionによる総合的評価の平均スコアは4.8(最小値:4、最大値:5.5)だった。

本論では、2008年4月から12月までに継続して行ったライティング課題を、任意の4時期に分類した。Term1は第1回目～6回目、Term2は第7回目～12回目、Term3は第13回目～18回目、Term4は第19回目～24回目である。それぞれの時期に書かれたエッセイを、第2言語習得(SLA)研究分野の中の、特にライティング研究で提唱されてきた評価指標、そしてコーパス言語学で提案されてきた言語情報処理方法論によって分析する。分析の目的は、「Term1からTerm4にわたって、向上が見られる学習者のライティングパフォーマンスは何か」ということを特定することである。

3. 継続したライティング活動の効果検証

3.1 分析の観点

Wolfe-Quintero, Inagaki & Kim (1998)⁹に従うと、ライティングの評価観点は、流暢さ(fluency)、文法的複雑さ(grammatical complexity)、語彙的複雑さ(lexical complexity)、正確さ(accuracy)の4つに大別される。一方Brown (2007)¹⁰では、内容(content)、構成(organization)、談話(discourse)、統語(syntax)、語彙(vocabulary)、書式(mechanics)の評価観点が提案されている。本論では、言語情報処理技術を利用すること、つまりコンピュータにより自動的に数値結果を求められることを重視し、以下の観点を採用する。

- (1) 流暢さ(fluency)
- (2) 統語的複雑さ(syntactic complexity)
- (3) 語彙的複雑さ(lexical complexity)
- (4) 内容の複雑さ(content)
- (5) 談話構造の複雑さ(discourse)

3.2 分析の方法

上述の各観点は、「コンピュータによって自動的に処理することができる」という点を重視し、以下の方法で分析を行った。

- (1) 流暢さ(fluency)
エッセイの総語数
- (2) 統語的複雑さ(syntactic complexity)
平均文長
(総語数/センテンス数)
接続詞の使用

(CLAWS⁷による接続詞タグの正規化頻度)
前置詞の使用

(CLAWSによる前置詞タグの正規化頻度)
不定詞の使用

(CLAWSによるto不定詞タグの正規化頻度)

(3) 語彙的複雑さ(lexical complexity)

語彙の豊富さ(lexical variation)

(Guiraud Index: Type / $\sqrt{\text{Token}}$)

語彙密度(lexical density)

(内容語÷総語数×100)

語彙の洗練度(lexical sophistication)

(JACET8000⁸によるLexical Frequency Profile⁹)

(4) 内容の複雑さ(content)

一般的・抽象的語彙の使用

(USAS¹⁰により“general and abstract terms”と判定された語の正規化頻度、例:do, gain, activity, idea, weird, important)

心的行動・状態・過程に関わる語彙の使用

(USASにより“psychological actions, states and processes”と判定された語の正規化頻度、例:learn, experience, conclude, able, think)

(5) 談話構造の複雑さ(discourse)

談話標識(discourse marker)の使用

次の語句の正規化頻度; so, because, and, but, or, ~ly speaking, in * case/s, in general, on the whole, for example/instance, etc. (Swan, 2005)⁶

各指標について、4つの時期(Term1～Term4)に収集されたデータ群の平均値をそれぞれ求める。それら平均値の間に有意な差が見られるか、一元配置の分散分析(ANOVA)により検定する。有意差が確認できた場合、その点において継続的なE-learning学習が効果をもたらしたと判断する。

3.3 分析の結果とその解釈

統計的な有意差が見られたのは、「流暢さ」と「語

⁷ 英語品詞タグ自動付与プログラム、英国ランカスター大学のコーパス研究グループによって開発された。
<http://ucrel.lancs.ac.uk/claws/>

⁸ 大学英語教育学会(JACET)のプロジェクトにより選定された日本人英語学習者がコミュニケーションを行う上で必要とされる8000語の語彙リスト(相澤他, 2005⁴)。

⁹ Laufer & Nation (1995)⁶による、語彙の洗練度(低頻度語の使用)を考慮した指標

¹⁰ 英単語意味タグ自動付与プログラム、英国ランカスター大学のコーパス研究グループによって開発された。
<http://ucrel.lancs.ac.uk/usas/>

彙の洗練度」の2項目だった。流暢さに関しては、Term 1 と Term 4 の総語数の平均値間（それぞれ 364.3 語と 430.1 語）に統計的有意差が認められた ($F=2.445$, $df=3$, $p=0.05$)。その推移は図4の通りで、比較的急に、そして継続的に書く量が増えていくのわかる。大学1年生にとってそれまで400語以上の英語エッセイを書いた経験が豊富にあったとは考えられず、総語数の増加は「エッセイライティングへの慣れ」の現れだと解釈できる。しかしながら、活動を行って7~8ヶ月経ったTerm 4において有意差が認められた、つまり長期間にわたり伸びが見られたことから、単に「慣れてきた」だけではなく、継続的活動が流暢さの向上に対して効果を発揮すると解釈することができるだろう。

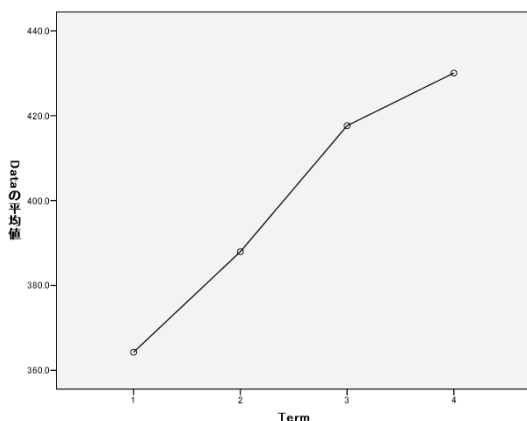


図4 Term 1 から Term 4 までの総語数の推移

語彙の洗練度に関しては、JACET Level 2 に属する単語の使用が Term 1 (4.76%) と Term 4 (6.01%) ($F=2.366$, $df=3$, $p=0.1$)、JACET Level 4 の単語の使用が Term 1 (1.47%) と Term 4 (1.85%) ($F=2.239$, $df=3$, $p=0.1$) を比較した際にそれぞれ増えていることが確認された。Term 4 になって使用するようになった具体的な単語は、以下表2 (Level 2 単語) と表3 (Level 4 単語) である。

表2 Term 4 に出現した JACET Level 2 単語

| 品詞 | 単語 |
|-----|--|
| 名詞 | childhood, coach, department, observation, path, possibilities, scores |
| 動詞 | connects, appreciated, invited, observing, preventing, refer, shot |
| 形容詞 | actual, wise, pure |
| 機能語 | nevertheless, further |

表3 Term 4 に出現した JACET Level 4 単語

| 品詞 | 単語 |
|----|-------------------------------|
| 名詞 | participants, tutor, outcomes |
| 動詞 | continued, evaluating |
| 副詞 | severely, currently |

表に見られるように、Term 4 になって使用するようになるのは 名詞、動詞、形容詞、副詞から成る内容語が中心になっている。機能語が文法的役割を果たすのに対し、内容語は意味を伝達する。つまり、内容語の豊富さは語彙力の充実さを表すだろう。当該の学生たちは比較的英語力が高いため、文法知識は活動開始当初からある程度は備わっており、ライティング活動を通じて培われた知識は語彙力だと推測できる。また、*Criterion* の診断的フィードバックのうち形式 (style) に関わる項目に「同一単語の繰り返し」(Repetition of Words) があり、この点を指摘される受講生が大勢いた。そのためこの点に注意し、なるべく同義語を用いるように対面授業においても指導をした。その結果がより洗練された単語の使用につながったのかもしれない。

4. おわりに

本論は、英語を専攻する比較的能力の高い学習者に対して行われている E-learning 学習法を紹介するとともに、継続的な E-learning による英語エッセイライティング活動が学習者の英語力向上に寄与するのか、するとすればどのライティングパフォーマンスが向上するのかを調査したものである。E-learning を通して得られた作文データは、学習者コーパスとして保存され、その分析には第2言語習得研究およびコーパス言語学研究で提案されてきた評価指標を用いた。その結果、「流暢さ」と「語彙の洗練」において活動の効果が認められた。

これら2つの指標以外の項目で統計的有意な差が見られなかった理由として考えられるのは、以下の2点である。

- (1) E-learning によるライティング練習は効果が得られない
- (2) 採用した評価指標が適切ではなかった

本論ではこれ以上の議論をすることはできないが、E-learning 学習は毎年授業で行われており、異なる年のデータを比較検討すること、ここで得られた結果の信頼性を検証することができるだろう。また、今後も第2言語ライティング研究およびコーパス言語学分野の知見を応用することで、作文データの分析を継続することができる。

このような調査を重ねることにより、授業で行わ

れている活動がはたして受講者のどのような能力を伸ばしているのかを知ることができる。言語習得研究の教育実践への適切な応用と言えるだろう。

謝辞

本研究の一部は、情報科学研究所研究助成によるものである。

参考文献

- (1) Monaghan, W. & Bridgeman, B. : E-rater as a Quality Control on Human Scores. R&D Connections, http://www.ets.org/Media/Research/pdf/RD_Connections2.pdf (2005.4)
- (2) Wolfe-Quintero, K., Inagaki, S. & Kim, H-Y. : Second Language Development in Writing: Measures of Fluency, Accuracy, & Complexity. Hawaii University Press. (1998.12)
- (3) Brown, H.D. : Teaching by Principles. Pearson ESL. (2007.4)
- (4) 相澤一美、石川慎一郎、村田年、磯達夫、上村俊彦、小川貴宏、清水伸一、杉森直樹、羽井佐昭彦、望月正道：『JACET8000 英単語「大学英語教育学会基本語リスト」に基づく』、桐原書店 (2005.11)
- (5) Laufer, B. & Nation, P. : “Vocabulary size and use: Lexical richness in L2 written production”. Applied Linguistics, 16, pp. 307-322 (1995.9)
- (6) Swan, M. : Practical English Usage. Oxford University Press. (2005.7)

(2011 年 9 月 30 日受付)
(2011 年 12 月 21 日採録)