

英文産出に与えるフィードバックの効果検証 —Criterion[®]とピア・フィードバックに焦点をあてて—

細越響子^{*1} 金丸敏幸^{*1} 高橋幸^{*1} 田地野彰^{*1}
京都大学^{*1}

(k.hosogoshi@hs5.ecs.kyoto-u.ac.jp, kanamaru@hi.h.kyoto-u.ac.jp,
takahashi.sachi.4n@kyoto-u.ac.jp, akira@tajino.mbox.media.kyoto-u.ac.jp)

1 はじめに

大学における英語教育の主要目的の一つは、学生の「学術目的の英語」(English for Academic Purposes : 以下, EAP)の運用能力の育成である。EAP の一技能であるアカデミック・ライティング(以下, AW)は、学問の場において必要なライティングの技能を指し、レポートの作成から論文執筆まで幅広いライティングの技能を含む。ライティング指導においては、執筆の過程で学生の草稿にコメントや添削など何らかの反応(フィードバック)を与えることが有効であるとされる(Leki, 1998)。しかしながら、日本の大学における AW 授業では 1 人の教師が比較的大人数のクラスを複数担当することも珍しくなく、毎回、妥当性の高い評価基準に従って教師 1 人でフィードバックを提供することは容易ではない。

そこで、われわれは、オンライン英文構成支援ツール Criterion[®]に着目し、その効果的な活用と教育効果について検討してきた(田地野他, 印刷中)。Criterion とは米国 Educational Testing Service により開発された英文自動評価ツールで、TOEFL[®] の自動採点システム e-rater[®]と連携している。エッセイを 1~6 点の 6 段階で 20 秒以内に採点し、加えてその誤りを、文法(Grammar)、語法(Usage)、構造(Mechanics)、文体(Style)、構成と発展(Organization & Development)の 5 つの観点から分析し指摘するものである。この結果は、教師によるフィードバックと中程度(Weigle, 2010)、あるいは、高い(Attali & Burstein, 2006)相関があると報告されている。

これまでの研究から、Criterion によるフィードバックはエッセイの質の向上に効果があり学生はその使用を好意的に捉えていること、一方、内容や構成面に関する指摘についてはその解釈に他者の支援を必要としていることが明らかになった(田地野他, 印刷中)。そこで本研究では、Criterion とピア・フィードバック(peer feedback : 以下, PF)の併用を試み、その教育効果を検証した。PF とは学生同士のフィード

バックを指し、学生の自己修正能力を養成し(Hyland & Hyland, 2006)、教師フィードバックと同様の効果を示すとの報告もある(Hedgcock & Lefkowitz, 1992)。

本研究の調査では、Criterion 使用学習において PF 実施群と非実施群を設定し、学習の前後で Criterion の総合スコアが上昇した学生を対象に、どのような言語的特徴の変化がエッセイの質の向上につながったか、また、フィードバックの方式によりその傾向に差異が見られたかを、Criterion の評価、語彙的特徴、統語的特徴、談話的特徴の 4 観点から分析した。

2 分析方法

2.1 調査時期と対象

本研究の調査は、2011 年 10 月に 2 週間にわたり実施された。参加者は、国立大学学部 1 年生の AW クラス受講者 4 クラス 103 名である。このうち両方の授業に出席し、かつ、書き直し後の総合スコアが 1 点以上向上した者のみ 74 名を分析対象とした。

2.2 調査内容と手順

本調査では、参加者に TOEFL 形式のエッセイ課題に取り組みさせた後、群ごとに指定された方式でフィードバックを利用させ、その結果をもとに書き直しをさせた。エッセイ課題は、「大学卒業後は、故郷で暮らすべきか否か」について立場を明確にして論じさせる説得型(persuasive)のものとした。

まず、第 1 週目に Criterion の使い方を説明した後、30 分間で第 1 稿を入力し提出させた。辞書は使用不可とした。授業後、第 1 稿の総合スコアに基づき参加者を PF 群(A・B クラス, 35 名)と非 PF 群(C・D クラス, 39 名)に分けた(表 1 参照)。

第 2 週目には、まず 20 分間で第 1 稿の Criterion によるフィードバックを各自確認させた。その後、PF 群のみ、2 人 1 組でエッセイを交換させ 30 分間後述の PF ガイドシートを使用した PF 活動を行わせた。最後に、再び全員に第 2 稿を Criterion 上に 20 分間で入力し提出させた。

2.3 PF ガイドシート

PF 群には、PF の際に着目すべき点を質問形式で示した全 7 項目の PF ガイドシートを配布した。内容は、先行調査の結果を踏まえ、内容面(論理の一貫性、明瞭性)や構成面(主題文や結論の有無)の確認を促すものとした。学生は、例えば「①このエッセイの主題文は何ですか?文中に下線を引き①と書いて下さい」などの問いに対し、該当する箇所に下線を引き適宜コメントを記入する形で PF を行った。

2.4 分析方法

Criterion と PF の教育効果を検証するため、どのような言語的特徴の変化がエッセイの質の向上につながったか、また、フィードバックの方式によりその傾向に違いが見られたかを、Criterion の評価、語彙的特徴、統語的特徴、談話的特徴の 4 観点から 2 要因の混合計画の分散分析で検定した。また、それぞれの効果量 η^2 も算出した。

(1) Criterion の評価

Criterion の評価に関しては、総合スコア(1~6 点)と文法、語法、構造の各エラー数、文体のコメント数の変化を検討した。

(2) 語彙的特徴

語彙的特徴としては、1 エッセイあたりの延べ語数と異なり語数を算出し比較した。同様に、British National Corpus (BNC)に基づき各語の語彙レベルを参照し、95%カバー率(Nation, 2001)に相当する単語の語彙レベルをそのエッセイの語彙レベルとした。分析には Web VP/BNC-20 を用いた。

(3) 統語的特徴

エッセイの統語的特徴を検討するために、まず

Sentence Extractor/T-Unit Calculator を使用し T-unit を求め、GENIA Tagger Demo を使用して S-node を計算した。T-unit と S-node は、それぞれ“An S-node is equivalent to a verb phrase (VP) (both finite and infinite) [...] a T-unit was defined as a main clause plus any subordinating clauses” (Ishikawa, 2006, p. 202)と定義されるもので、統語的特徴を検討する際に使用される指標である。

統語的の流暢さについては、Words/T-unit(T-unit あたりの延べ語数)を指標として採用し、統語的複雑さは S-nodes/T-unit(T-unit あたりの動詞句数)を求めることで分析した。

(4) 談話的特徴

最後に、談話的特徴は Criterion の構成と発展に関する指摘を参照し、序文(Introduction)、主題(Thesis)、趣旨(Main Ideas)、展開文(Supporting Ideas)、結論(Conclusion)の各要素に相当すると判定された T-unit の数をそれぞれ算出して比較した。

3 分析結果及び考察

3.1 Criterion の評価

以下の表 1 は、Criterion による評価の比較結果である。表 1 より、書き直しによって、文法、語法、構造に関するエラーが有意に減少することがわかった。一方、その効果に PF の有無による影響はみられなかった。この結果は、Criterion により適切な語句の使用や文法的機能の選択、正書法に関する指摘が与えられることで、学生はエッセイの正確性を高めることができることを示唆している。

表 1 Criterion の評価の比較

	PF 群 ^a		NPF 群 ^b		主効果 ^c		交互作用 ^d
	第 1 稿	第 2 稿	第 1 稿	第 2 稿	PF の有無	書き直し	PF × 稿
総合(Holistic)	3.11	4.26	3.31	4.46	2.00	616.54**	0.01
スコア	(0.57)	(0.55)	(0.69)	(0.67)	(.01)	(.45)	(.00)
文法(Grammar)	1.31	1.06	1.31	0.90	0.14	5.57*	0.29
エラー数	(1.04)	(1.01)	(1.18)	(1.22)	(.00)	(.02)	(.00)
語法(Usage)	2.71	2.06	2.97	1.44	0.33	19.31**	3.11
エラー数	(1.83)	(1.58)	(1.95)	(1.37)	(.00)	(.09)	(.02)
構造(Mechanics)	4.29	2.31	4.44	1.87	0.06	55.47**	0.95
エラー数	(3.35)	(2.85)	(2.86)	(2.13)	(.00)	(.14)	(.00)
文体(Style)	35.80	33.86	35.05	34.38	0.00	0.78	0.19
コメント数	(12.49)	(17.39)	(11.60)	(15.21)	(.00)	(.00)	(.00)

Note. ** $p < .01$, * $p < .05$; ^a $n = 35$, ^b $n = 39$; ^{a,b} 上段= M , 下段= SD ; ^{c,d} 上段= F 値, 下段= 効果量(η^2)

表2 語彙的特徴の比較

	PF 群 ^a		NPF 群 ^b		主効果 ^c		交互作用 ^d
	第1稿	第2稿	第1稿	第2稿	PFの有無	書き直し	PF × 稿
延べ語数	172.71	226.23	182.26	229.23	0.44	172.09**	0.73
(Token)	(28.39)	(42.41)	(44.53)	(52.21)	(.00)	(.25)	(.00)
異なり語数	85.46	105.66	88.44	105.10	0.11	229.12**	2.10
(Type)	(13.14)	(17.49)	(15.35)	(17.51)	(.00)	(.25)	(.00)
語彙レベル ^e	1.60	1.66	1.69	1.72	0.46	0.71	0.10
	(0.49)	(0.53)	(0.56)	(0.50)	(.00)	(.00)	(.00)

Note. ** $p < .01$, * $p < .05$; ^a $n = 35$, ^b $n = 39$; ^{a, b}上段 = M , 下段 = SD ; ^{c, d}上段 = F 値, 下段 = 効果量(η^2)

^e単位 = 千語, ワードファミリー換算

一方, 文体については書き直しの前後で有意な変化が見られなかった. 文体の指摘には同一語の多用や1文の長さに関するものが含まれるが, 特に学生はキーワードや“ I ”などの主体表現を多用せずに自分の主張を展開することに困難を示しており, この点については, ピアの助言だけではなく教師など熟達者の支援が必要となるであろう.

3.2 語彙的特徴

次に, 語彙的特徴については, 表2に示すように Criterion 使用の結果, 全体の使用語数が平均して50語近く増加していた. これは, 議論を補強するため追加情報を挿入したためであると考えられる. しかしながら, PFの有無による影響は観測されなかった.

一方, 使用語彙のレベルは書き直しの前後で変化がみられなかった. これは, エッセイ課題が一般的事象に関するものであり, その議論は使用頻度の高い語彙のみで充足できたためと考えられる.

3.3 統語的特徴

統語的特徴の変化は, 他の観点とは様相が異なっていた. 表3に要約したように, 群の違いに関わらず書き直しの結果1エッセイあたりの文の数は平均して3文ほど増加していたが, 1文あたりの統語的

流暢さ(Words/T-unit)と統語的複雑さ(S-nodes/T-unit)は有意に減少していた. このことから, 学生の1文ごとの修正は, 元の文に新たな要素を付加するよりも, 文意がより明瞭になるよう構造を簡潔にしようとする傾向があったと推察される.

3.4 談話的特徴

最後に, Criterionの指摘に基づき全体的な談話構造の変化を比較したのが表4である. 表4から, Criterionを使用した結果, 特に談話の骨格をなす序論と結論, 議論を展開する要素である趣旨と展開文の数が有意に増加することがわかった. 一方, PF活動はとりわけ内容や構成面に気づきを促すことを意図したものであったが, Criterionのみのフィードバックと比較して有意な効果は観察されなかった.

以上の結果と, エッセイあたりの文の総数は増加したものの統語的流暢さや複雑さは減少したという前項の結果から, 内容や構成面に関する学生の修正方法は, 論理展開が不明瞭であったり論拠が不十分だと指摘された部分について, 既存の文の構造を変えて補足情報を追加するよりは, 新たに文を挿入することで不足情報を補完しようとするものであったと考えられる.

表3 統語的特徴の比較

	PF 群 ^a		NPF 群 ^b		主効果 ^c		交互作用 ^d
	第1稿	第2稿	第1稿	第2稿	PFの有無	書き直し	PF × 稿
T-unit 数	11.17	14.71	11.33	14.64	0.00	111.10**	0.13
	(2.62)	(2.93)	(3.20)	(3.21)	(.00)	(.24)	(.00)
統語的流暢さ	15.88	15.53	16.53	15.80	0.61	5.58*	0.73
(Words/T-unit)	(2.52)	(2.22)	(2.33)	(2.41)	(.01)	(.01)	(.00)
統語的複雑さ	2.66	2.58	2.76	2.64	0.68	5.70*	0.28
(S-nodes/T-unit)	(0.44)	(0.35)	(0.57)	(0.42)	(.01)	(.01)	(.00)

Note. ** $p < .01$, * $p < .05$; ^a $n = 35$, ^b $n = 39$; ^{a, b}上段 = M , 下段 = SD ; ^{c, d}上段 = F 値, 下段 = 効果量(η^2)

表4 談話的特徴の比較

	PF 群 ^a		NPF 群 ^b		主効果 ^c		交互作用 ^d
	第1稿	第2稿	第1稿	第2稿	PFの有無	書き直し	PF × 稿
序論 (Introduction)	0.77 (0.93)	1.26 (1.10)	1.10 (1.03)	1.38 (1.21)	0.99 (.01)	12.91** (.03)	0.91 (.00)
主題 (Thesis)	2.03 (1.13)	1.77 (1.35)	1.56 (1.32)	1.23 (1.19)	3.65 (.04)	4.86* (.01)	0.08 (.00)
趣旨 (Main Ideas)	1.54 (0.55)	2.26 (0.50)	1.59 (0.59)	2.08 (0.62)	0.42 (.00)	49.30** (.22)	1.76 (.01)
展開文 (Supporting Ideas)	5.49 (1.79)	7.34 (2.00)	5.46 (2.31)	7.74 (2.77)	0.16 (.00)	65.73** (.17)	0.69 (.00)
結論 (Conclusion)	1.03 (1.11)	1.57 (1.00)	1.03 (1.07)	1.56 (1.22)	0.00 (.00)	14.33** (.06)	0.00 (.00)
その他 (Others)	0.46 (0.50)	0.57 (0.50)	0.69 (0.46)	0.72 (0.45)	3.43 (.04)	2.47 (.01)	0.99 (.00)

Note. ** $p < .01$, * $p < .05$; ^a $n = 35$, ^b $n = 39$; ^{a,b} 上段 = M , 下段 = SD ; ^{c,d} 上段 = F 値, 下段 = 効果量(η^2)

4 おわりに

本研究から、Criterion によるフィードバックは、語彙や文法使用の正確性や正書法、総語数と総文数、そして論理展開の明確な文章構成といった項目の修正を促し、英文産出の質的な向上につながる事が明らかとなった。具体的には、学生は Criterion の指摘に従い語句の使用や文法に関する誤りを訂正するとともに、論理展開が不明瞭または不十分と判定された箇所については新たに文を挿入し補足することでエッセイの質の向上を図る傾向があることが示唆された。

また本研究では、内容や構成面の修正に貢献することを期待して Criterion 使用学習に PF という異なる形式のフィードバックを併用したが、今回の結果からは PF の書き直しに対する効果は観察されなかった。今回は PF 活動に際し、ガイドシートを除き事前指導を行わなかったため、学生は目標とすべきライティングの前提知識が不十分であった可能性がある。ゆえに、AW 授業において PF を活用するためには、事前に教師主導で PF の方法を学習する機会を設け、論理的な議論の展開や適切な表現のモデルを理解しておくことが重要であると考えられる。

今後は、他のレベルの学生のエッセイを分析し、Criterion やピアによるフィードバックが有効な要素と、学習に際して教師の介入が求められる要素を明らかにすることで、AW 授業におけるフィードバックの在り方をさらに検討する予定である。

注

本研究は、科学研究費補助金(基盤研究(B) 課題番号 22320102)の助成を受けて実施された。

参考文献

- Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater[®] V. 2. *The Journal of Technology, Learning, and Assessment*, 4(3), 1-30.
- Hedgcock, J., & Lefkowitz, N. (1992). Collaborative oral/aural revision in foreign language writing instruction. *Journal of Second Language Writing*, 1(3), 255-276.
- Hyland, K., & Hyland, F. (2006). State of the art article: Feedback on second language students' writing. *Language Teaching*, 39(2), 83-101.
- Ishikawa, T. (2006). The effects of task complexity and language proficiency on task-based language performance. *The Journal of Asia TEFL*, 3, pp. 193-225.
- Leki, I. (1998). *Academic writing: Exploring processes and strategies*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- 田田野彰, 細越響子, 川西慧, 日高佑郁, 高橋幸, 金丸敏幸(印刷中). “アカデミックライティング授業におけるフィードバックの研究—Criterion[®]を導入した授業実践からの示唆—.” 京都大学高等教育研究, 17.
- Weigle, S. C. (2010). Validation of automated scores of TOEFL iBT tasks against non-test indicators of writing ability. *Language Testing*, 27(3), 335-353.