

日本人大学生を対象とした統合型ライティングプロダクトの 自己評価と教員評価の一致度合い

木 幡 隆 宏

Differences between Japanese University Students' Self-Evaluation and Teacher Evaluation on a Reading into Writing Task

Takahiro KOWATA

1. はじめに

英語技能は「読む」、「聞く」、「書く」、「話す」の4技能の観点から評価されることが多いが、Common European Framework of References for Languages (CEFR) では「話す」の技能を一方的な発話を指す「発表」と相手との対話を指す「やり取り」の2つの技能に分けて5つの技能として扱っており (Council of Europe, 2001), この考え方が広まってきている。日本でもこれを4技能5領域として扱うことが、小・中・高等学校の次期学習指導要領で組み込まれることになっている (中央教育審議会, 2016)。また、それぞれ単一の技能ではなく「読んだことについて書く」や「聞いたことについて話す」、「読んだり聞いたりことについて書く」など、統合的な技能が求められてきている。

学習段階において、学習者が自らの能力を把握し次の学習につなげることは自律的な学習者として重要な能力である。しかし、発表技能においては自分の考えが相手に伝わるかどうかの判断が難しく、特に書く能力においては読み手の反応が即座にないために、より一層困難である。本研究では、学習者が読んだことについて書く統合型タスク (以下 RW タスク) において、日本人大学生英語学習者の自己評価が教員の評価とどの程度一致しているのかについて明らかにし、どの観点での自己評価のための指導が必要であるかについて検討する。

2. 研究の背景

松井 (2015) は、日本人の高校1年生を対象に、与えられたトピックについて書かせたライティングプロダクトを用いて、自己評価力を育成するための評価ルーブリック活用方法について実践を行った。その結果、内容面と構成面については生徒自身でのモニタリングが見込めるが、言語面 (語彙・文法) では困難であることが報告されており、評価観点によっては自己評価の難易度に差があることが分かっている。

長阪 (2003) は、日本人大学生を対象に、与えられたトピックについて書かせた2種類のライティングプロダクトを用いて自己評価と教師評価の相関を調べた。ESL Composition Profile (Jacobs, *et al.*, 1981) を用いた評価であったが、1種類のプロダクトでは内容、構成、語彙、語法で .40 ~ .54 の中程度の相関が見られた、もう1種のプロダクトでは語法のみ .32 の相関が見られた。

これらの研究から、評価観点によって教師と学習者の評価結果にばらつきがあり、学習段階やライティングプロダクトによってもそのばらつき方が異なる可能性があることが分かった。しかし、RW タスクの自己評価研究は十分になされておらず、また使用する採点基準によってはさらに異なる傾向が見られる可能性がある。

書くことを最終プロダクトとした統合技能テストとしては、Educational Testing Service (ETS) が2005年に導入したTOEFL iBTに、聞いた内

容と読んだ内容を元にして書く能力を測定する問題が含まれている。このように世界では早くから統合的なライティング技能を測定したテストが普及しているが、日本国内の主要な大規模標準テストである英検やGTECなどにはこの種の問題が含まれておらず、RWタスクが含まれるものは大学入試のために開発され2014年から開始されたTEAPおよびその後開発されたTEAP CBTに限られている。

日本人英語学習者は高校時点ではあまり統合型作文の指導を受けておらず(木幡, 2015), 大学生の授業ではRWタスクを用いた統合的なライティング能力の育成が求められている。その評価手段として大規模標準テストの評価基準を参考にすることが考えられるが、TOEFL iBTは多くの大学生にとっては難易度が高く、TEAPの評価基準はCERFのA2未満からA2, B1, B2のレベル幅の広い4段階評価であるため、比較的近い能力を持った学習者が集まる教室環境での使用には適さない。一方、Trinity College Londonで開発したテストIntegrated Skills of EnglishはCEFRのレベルごとに問題が作成され、それぞれのレベルに評価基準が設定されているため、日本人大学生が教室環境で使用する目的にはより一致する。

3. 調査

3.1 目的

日本人大学生英語学習者の書いたRWタスクのライティングプロダクトについて、学習者の自己評価と教員評価の一致度合いを調べ、自己評価が一致しやすい観点と一致しにくい観点を特定する。

3.2 参加者

国立大学保健学科の2年生30名が調査に参加した。すべての学生は同じ授業を受講していた。また授業を担当する教員1名が教員評価者として調査に参加した。

3.3 調査方法

まず、参加者は授業中にRWタスクに取り組み、その次の授業時に指定された採点基準を用いて自己採点を行った。また、授業担当教員が同じ採点基準を用いて採点を行った。

使用したタスクの指示文は以下の通りで、Tipping in the United Statesというタイトル英語の文章を読み、その内容を用いながらShould all restaurants in the United States introduce anti-tipping policies?というお題に対して、国際交流のためのウェブサイトに関心のある自分の意見の記事を100～130語で書くものであった。また、参加者は辞書等のツールを使用できず、制限時間は60分だった。なお、Tipping in the United Statesの文章は2017年度第1回英語検定試験2級の問題から引用した。

<タスクの指示文>

Write an article for an intercultural website. On this site, university students all over the world exchange their unique ideas. Use the information from the article "Tipping in the United States" to write your opinion about the topic below.

"Should all restaurants in the United States introduce anti-tipping policies?"

Write your article of 100–130 words. Try to use your own words as far as possible – don't just copy sentences from the reading text. Those who do not read the text also read your article.

採点基準は、ISEI Reading-into-Writing Rubricを参考としながら、「タスクの達成」、「文章構成」、「テキスト情報の利用」、「正確さ」の4観点からなる分析的評価基準を日本語で作成した(資料参照)。ISEIはTrinity College Londonで開発したIntegrated Skills of Englishのレベルのひとつであり、CEFRのB1に相当するレベルということから受験者の能力帯と一致する。また問題の内容、語数もほぼ一致することから参考にすることが妥当であると判断した。全く同じものを使用しなかった理由としては、RWタスクが同一でないため不要な記述を削除したことと、参加者が英語の記述の理解に苦慮することが挙げられる。各観点の採点基準は0から4の5段階得点とした。なお、評価にあたっては各得点記述文のすべての要素が満たされた最も高い値をとった。

参加者の自己評価にあたっては、教員が採点基準について説明した上で自己評価を行った。その後教員が同じ評価基準を用いて評価を行った。

3.4 分析

まず、学習者の自己評価と教員評価について記述統計を算出し、対応のある t 検定を行った。次に、4つの観点ごとに学習者の自己評価と教員評価のピアソン積率相関係数を算出した。

4. 結果と考察

表1は、学習者の自己評価と教員評価の記述統計を示している。 t 検定の結果、「文章構成」と「正確さ」の観点では有意な差が見られなかった ($t = .311, df = 29, ns; t = -1.651, df = 29, ns$)。しかし、「タスクの達成」の観点で教員の方が有意に高い評価を行い ($t = 3.440, df = 29, p < .05$)、テキスト利用の観点で学習者の方が有意に高い評価を行っていた ($t = -2.164, df = 29, p < .05$)。

表1. 自己評価と教員評価の記述統計 (N = 30)

		タスク 達成	文 章 構 成	テキスト 情報の利用	正確さ
自己評価	M	3.03	2.63	2.70	2.80
	(SD)	(.81)	(.81)	(.70)	(.81)
教員評価	M	3.70	2.70	2.33	2.57
	(SD)	(.70)	(.88)	(.92)	(.82)

表2は、学習者の自己評価と教員評価の相関係数を示している。「テキストの利用」の観点は.373、「正確さの観点」は.545の有意な相関があった。しかし、「タスクの達成」と「文章構成」の観点では相関は有意でなかった。

表2. 自己評価と教員評価の相関

評価観点	タスク 達成	文 章 構 成	テキスト 情報の利用	正確さ
Pearson r	.018	.034	.373*	.545**

* $p < .01$; ** $p < .001$

4つの観点のうち、「正確さ」は自己評価と教員評価に有意な差がなく、中程度の相関が見られたことから、最も学習者の自己評価と教員評価が一致していた。これは、長阪(2003)の語彙・語法と似た観点であり、同程度の一致度を示していたと解釈できる。また、日本人大学生英語学習者にとって「正確さ」はより正確に自己評価をすることができる観点であり、この点では松井(2015)で示されている高校1年生段階よりも自己分析力

が高くなっているのではないかと推測できる。

「テキストの利用」の観点では、自己評価と教員評価に低い相関が見られたが、学習者が教員よりも有意に高い評価を行っていた。これは、テキスト内容を自分のことばで表現するという点については互いの評価が一致しやすいが、その情報を意見と効果的に関連させているか否かを判断することがより難しく、学習者の方がよりあまくその関連性を評価していると考えられる。

「タスクの達成」の観点では、教員評価が自己評価より有意に高かったが、その理由としては参加者が「タスクのお題に沿った明確な自分の意見を述べている」という記述文の求める内容を高くとらえてしまったことが考えられる。また、このずれが有意な相関が見られなかった要因の大きなひとつであるが、同時に教員評価に天井効果が認められた点も挙げられる。

「文章構成」の観点では、自己評価と教員評価に有意な差がなかったにも関わらず、有意な相関は見られなかった。両者の採点結果をひとつずつ比較すると、評価が一致した参加者が13名、教員よりも高く評価した参加者が9名、教員よりも低く評価した参加者が8名と、4観点の中で最も一致度が低かった。この点については様々な理由が考えられ、例えば英語の正確さも関連してくる可能性がある。学習者本人は書いた意味を理解しているのに対して、読み手の教師には伝わっていない部分があるため、論理展開の認識にずれが生じてしまう。その場合は学習者の方がより高い評価をする可能性がある。また学習者が自身の書いた作文がどの程度論理的であるか客観的に判断することが難しいと感じ、実際よりも低く評価した可能性もある。

5. 結論

日本人大学生英語学習者は、RWタスクにおいて「タスクの達成」、「文章構成」、「テキスト情報の利用」、「正確さ」の4観点のうち、「正確さ」については教員評価とより一致した自己評価を行うことができたが、その他の3観点については改善の余地があった。特に「文章構成」の自己評価傾向はばらばらであったため、より集中的な指導が必要とされる。また、統合型技能の特徴である情報利用の適切さという点でも自己評価の改善が

見込めるため、より効果的な情報の使用についての指導が必要である。

6. 今後の課題

本研究に参加した学習者は30名と少なく、また使用したタスクも1種類であることから、より多くの学習者を対象とし、他のタスクでも同様の結果が出るかどうかの検証が必要である。また、教員の評価者が1名であることから、評価が正確でない可能性がある。そのため、今後の研究では評価者を複数人にするこゝで評価の信頼性を確保する必要がある。

7. 謝辞

本研究はJSPS 科研費 17K13493 の助成を受けたものです。

参考文献

木幡隆宏. (2015). 『大学入試における自由作文問題の学習と指導への波及効果』 博士論文 (学術) 東京外国語大学.

中央教育審議会. (2016). 『幼稚園, 小学校, 中学校, 高等学校及び特別支援学校の学習指導要領等の改善及び必要な方策等について (答申)』 (http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/_icsFiles/afieldfile/2017/01/10/1380902_0.pdf)

長阪朱美. (2003). 「学習者の英語ライティング能力の認識」『恵泉女学園大学人文学部紀要』 15, 95-111.

松井市子. (2015). 「評価 rubric を活用した英語ライティング力と自己評価力の育成をねらった実践」『EIKEN BULLETIN』 27, 146-164.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Jacobs, H. L., Zinkgraf, S.A., Wormuth, D.R., Hartfiel, V.F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.

Trinity College London. (2016.) *Integrated skills in English (ISE) Specifications – Reading & Writing*. (www.trinitycollege.com/resource/?id=6299)

<資料>

資料 使用した評価基準 タスクの達成

4	タスクのお題に沿った明確な自分の意見を述べている。 提示されたテキストの情報を使用している。 要求された語数を満たしている。 要求されたライティングのスタイルを満たしている。
3	タスクのお題に沿った明確な自分の意見を述べている。 提示されたテキストの情報を使用している。 要求された語数をおおよそ満たしている (90 ~ 140 語)。 要求されたライティングのスタイルを満たしている。
2	タスクのトピックに沿った自分の意見を述べている。 提示されたテキストの情報を使用している。 要求された語数がある程度満たしている (80 ~ 150 語)。
1	タスクのトピックに沿った自分の意見を述べている、または提示されたテキストの情報を使用している。 要求された語数をほとんど満たしていない (80 語未満, 151 語以上)。
0	タスクの要求を全く満たしていない。

文章構成

4	詳細を明確に示しながら，論理的に展開している。 接続表現や指示詞を正確に使用している。
3	多くの部分で詳細を明確に示しながら，論理的に展開している。 多くの部分で接続表現や指示詞を正確に使用している。
2	ある程度詳細を明確に示しながら，論理的に展開している。 ある程度，接続表現や指示詞を正確に使用している。
1	展開が論理的ではない部分が多く，詳細を示している部分が少ない。 接続表現や指示詞の使用が正確さに欠けている。
0	採点不可能。

テキスト情報の利用

4	テキストの主旨を理解し，自分の意見と効果的に関連させて情報を使用している。 テキスト内容を適切に自分のことばで表現しており，コピーしている部分は文章全体の1割以下である。
3	テキスト内容の大部分を理解し，自分の意見と十分関連させて情報を使用している。 テキスト内容を正確に自分のことばで表現できている部分が多く，コピーしている部分は文章全体の2割以下である。
2	テキスト内容の主要な部分を理解し，テキスト情報と自分の意見とを最低限関連させている。 テキスト内容を自分のことばで表現しようとしており，コピーしている部分は文章全体の3割以下である。
1	テキスト情報と自分の意見とのつながりが希薄である。 テキスト内容のコピーが目立ち，文章全体の3割を超える。
0	テキストの情報を使用していない。

正確さ

4	若干の細かなミスを除き，文法，語彙を正確かつ適切に使用している。 設定されている読み手にとって理解を妨げる誤りはない。
3	文法，語彙をほとんどの部分で正確かつ適切に使用している。 設定されている読み手にとって理解を妨げる誤りは1割以下である。
2	文法，語彙を多くの部分で正確に使用している。 設定されている読み手にとって理解を妨げる誤りは2割以下である。
1	文法，語彙の誤った使用が多い。 設定されている読み手にとって理解を妨げる誤りは3割を以下である。
0	設定されている読み手にとって理解を妨げる誤りは3割を超える。