

ChatGPT を活用した IELTS エッセイライティング学習の検討

Examination of IELTS Essay Writing Learning Using ChatGPT

外山愛華*1, 谷田貝雅典*2, 米谷雄介*3, 永岡慶三*4, 卯木輝彦*1
Aika SOTOYAMA*1, Masanori YATAGAI*2, Yusuke KOMETANI*3, Keizo NAGAOKA*4, Teruhiko UNOKI*2

*1関西外国語大学外国語学部

*1College of Foreign Studies, Kansai Gaidai University,

*2共立女子大学文芸学部

*2Faculty of Arts and Letters, Kyoritsu Women's University

*3香川大学創造工学部

*3Faculty of Creative Engineering and Design, Kagawa University

*4早稲田大学人間科学学術院

*4Faculty of Human Sciences, Waseda University

Email: s222119@kansai.gaidai.jp, unoki@kansai.gaidai.ac.jp

あらまし: 本稿では、外国語学習者にとって効果的な ChatGPT の活用方法を明らかにすることを目的に、IELTS ライティングタスク 2 を対象に、ChatGPT が学習者のエッセイに対して適切な評価および採点を行うことができるか、IELTS の運営団体が公開している練習問題の解答例をもとに検証を行った。その結果、ChatGPT による採点結果は、公認の採点者による結果と有意な差があるとは認められず、概ね妥当であると考えられることが分かった。

キーワード: 生成系 AI, ChatGPT, 英語学習, エッセイライティング, IELTS

1. はじめに

ChatGPT などの生成系 AI を学習に積極的に活用することを前提とした議論が各所で盛んに行われている⁽¹⁾。特に語学学習では、語彙の強化や文法等の疑問の解決にとどまらず、その背後にある文化的背景についても ChatGPT により効果的に学ぶことができると期待されている⁽²⁾。

筆者らの先行研究^(3,4)では、ChatGPT を練習相手に特定のシチュエーションを想定して英会話を行うロールプレイ学習や、ChatGPT に対して英文エッセイの添削や改善提案を指示して自身が執筆したエッセイをより洗練された表現に仕上げていくエッセイライティングの学習において、学習者一人でも効率的な学習ができることが示された。

ChatGPT が語学学習の補助ツールとして非常に価値がある一方、その出力は文脈を理解したものではなく、大量のテキストデータから学習をしたモデルに基づき生成されるため、正確さや信頼性が必ずしも保証されているわけではない。

本稿では、IELTS (International English Language Testing System) ライティングタスク 2 を対象に、ChatGPT が学習者のエッセイに対して適切な評価、採点ができるかについて検証する。

2. 方法

2.1 IELTS ライティングタスク 2

IELTS は、英語圏の国々に留学、就労または移住をするための英語能力を測定する試験である。ライティングタスク 2 は、社会、教育、技術、環境など、幅広いトピックについて 250 語以上のエッセイを書

くことが求められる。採点は、英語教育の分野での経験と専門知識を持つ訓練された複数の採点者により行われ、4 つの観点「Task Response (課題への対応)」「Coherence and Cohesion (一貫性と結束性)」「Lexical Resource (語彙力)」「Grammatical Range and Accuracy (文法の範囲と正確さ)」ごとに、バンドディスクリプターと呼ばれる評価基準が明確に定義されており、1 から 9 まで 0.5 刻みの値でスコアが付けられる。

2.2 実験条件

IELTS の運営団体が公開している練習問題⁽⁵⁾から公認の採点者による採点結果付きの解答例 6 件を選択した。選択した解答例には、公開されている解答例のうち最低スコアであるスコア 4 の解答、および最高スコアであるスコア 8 の解答を含む。選択した解答例のテーマとスコアを表 1 に示す。観点別の評価結果は公開されていない。

表 1 実験に使用したサンプル解答

No.	テーマ	スコア
A	高齢者の介護費用分担	5
B	将来の健康水準	5.5
C	将来の健康水準	7.5
D	レジャーとしてのショッピング	6
E	家庭環境が子供に及ぼす影響	4
F	レジャーとしてのショッピング	8

ChatGPT に対して、選択した解答例およびその問題文を次のプロンプトとともに入力をする。

プロンプト: 「次のエッセイを IELTS タスク 2 の評価基準に従って評価してください。観点別お

よび総合的なバンドスコアを推定してください」プロンプトは日本語で入力する。バンドディスクリプターなど IELTS に関する追加の情報は一切入力しない。また、ChatGPT は、確率的なモデルであり、同じプロンプトを入力しても毎回異なる結果を出力するため、評価結果のばらつきを考慮し、ChatGPT によるスコアは各解答例に対して5回の試行を行った結果の平均値とする。

3. 結果と考察

公認の採点者によるスコアと ChatGPT によるスコアを図1に示す。両者の差の最大値は1.3、最小値は0、平均値は0.45である。バンドスコアの刻み幅が0.5であることを考慮すれば、採点者によるスコアと ChatGPT によるスコアの差は十分に小さいと考えられる。

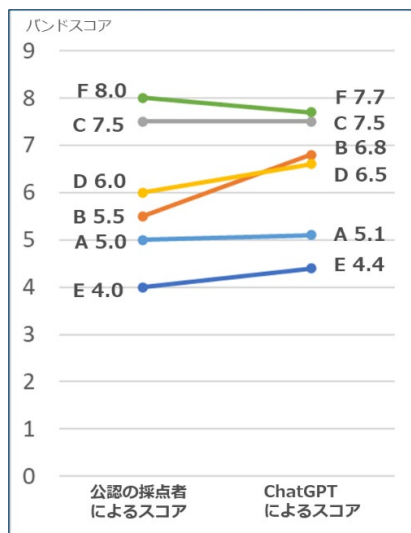


図1 採点者と ChatGPT によるスコアの比較

解答例 B について、公認の採点者によるスコアが5.5であるのに対し ChatGPT によるスコアは6.8と、その差は比較的大きい。公認の採点者の評価コメントを確認すると「段落分けは必ずしも論理的とは言えない」「スペル、単語の選択にいくつかの誤りがある」など内容に関する指摘に加え、「長さ不足」を指摘している。解答例 B は、250 ワード以上との規定に対し 198 ワードしか書かれていない。他方、ChatGPT による評価コメントでは「パラグラフの分け方やつながり」「スペルミスや単語の誤用」など公認の採点者と同等の指摘が減点項目としてあげられ、その理由や改善案が具体的に示されている。しかしながら、長さ不足についての指摘がない。ChatGPT は、バンドディスクリプターに記載されている評価基準については十分に把握をしており適切な指摘をしていたが、文字数制限に関する確認が抜けていた。これにより文字数の少ない解答に対しては高いスコアとしている可能性が考えられる。

最もスコアが低い解答例 E に対して、公認の採点

者は「わかりにくい」「不明瞭」「単語の選択とコロケーションの誤りが頻繁に見られる」など多くの指摘をしている。しかしながら、それがどの部分に対する指摘なのかを具体的に読み取れる記述はない。他方、ChatGPT は「一貫性がない」「不完全な文」「誤った単語の使用」「冠詞の誤り」など、公認の採点者とほぼ同様の指摘をしている。加えて、指摘箇所や修正案が具体的に示されていることが特徴である。

公認の採点者による採点と ChatGPT による採点の一貫性を評価するため、Wilcoxon の符号順位検定を行なった結果、両者の間には統計的に有意な差が見られなかった ($p = 0.18$)。これは、公認の採点者と ChatGPT が同じテストに対して本質的に異なる点数をつけているとは言えないことを示唆している。ただし、この研究では対象とした解答 ($n=6$) の数が少ないため、結果の一般化には注意が必要である。

以上のことから、文字数制限の確認など ChatGPT には一部不十分な面があるものの、ChatGPT による IELTS ライティングタスク 2 の解答例に対する評価は概ね妥当であると考えられる。

4. おわりに

本論文では、IELTS ライティングタスク 2 を対象に、ChatGPT が学習者のエッセイに対して適切な評価および採点ができるか、IELTS の運営団体が公開している練習問題の解答例をもとに検証を行った。その結果、ChatGPT による採点結果は、公認の採点者による結果と有意な差があるとは認められず、概ね妥当であると考えられることが分かった。また、ChatGPT では、バンドディスクリプターに従った観点別評価の結果を提示してくれることや、評価の低い箇所を具体的に示し改善案を提案してくれるなど、学習者にとってより有効な活用方法が期待できる。

今後は、エッセイ以外も含むさまざまなライティングタスクに対し、実践的な利用を通して ChatGPT の効果的な活用方法を確立していきたい。

本研究は JSPS 科研費 JP23K02646 の助成を受けたものである。

参考文献

- (1) 東京大学: "シンポジウム「未来の教育環境を創る: 生成系 AI への対応と展望」" (2023)
- (2) 金丸敏幸: "ChatGPT が語学(英語)教育に与えるインパクト," 第63回「大学等におけるオンライン教育とデジタル変革に関するサイバーシンポジウム」(2023)
- (3) 外山愛華, 折橋暖人, 谷田貝雅典, 卯木輝彦: "英語ライティング学習に対する ChatGPT 活用の検討," 学習分析学会 2023 年度第1回研究会, 東京 (2023)
- (4) 折橋暖人, 外山愛華, 谷田貝雅典, 卯木輝彦: "AI アバターに ChatGPT を活用した VR 外国語会話学習システムの検討," 学習分析学会 2023 年度第1回研究会, 東京 (2023)
- (5) IDP Education : "IELTS test preparation materials," <https://ieltsjp.com/japan/prepare/ielts-test-preparation-material> (参照 2023.06.05)