

生成 AI を用いた英語教材の可読性向上による多読支援

Supporting Extensive Reading by Improving the Readability of English Learning Materials Using Generative AI

寺尾 透哉^{*1}, 戴 憶菱^{*2}Toya TERAO^{*1}, Yiling DAI^{*2}^{*1}広島大学 情報科学部^{*1}School of Informatics and Data Science, Hiroshima University^{*2}広島大学大学院 先進理工系科学研究科^{*2}Graduate School of Advanced Science and Engineering, Hiroshima University

Email: b220522@hiroshima-u.ac.jp

あらまし：本研究では、生成 AI を用いて英語教材の可読性を学習者の読解レベルに応じて調整する多読支援手法を提案する。可読性指標に基づく数値的制御に加え、語彙や構文などの質的要因を考慮したプロンプト設計を行った。さらに、Web ベースの読書演習システムを構築し、行動ログおよび主観評価を用いた評価実験を実施した。その結果、生成 AI による可読性調整が、教材の可読性および学習者の没入感を向上させる可能性が示唆された。

キーワード：多読, 生成 AI, 可読性向上, 読書体験, フロー体験

1. はじめに

多読は、学習者が自身の能力に応じた易しい英文を大量に読むことで⁽¹⁾、語彙力や読解力の向上等に効果がある学習法である⁽²⁾。多読を効果的に行うためには、学習者の読解レベルに適合した教材を継続的に提供することが重要であるが、教材の入手と選定等の負担は大きく、導入上の課題となっている⁽³⁾。

近年、大規模言語モデルを中心とする生成 AI の発展により、文章の書き換えや難易度調整が可能となり、生成 AI を用いた教材生成の試みが報告されている⁽⁴⁾⁽⁵⁾。しかし、先行研究の多くは生成結果の可読性指標評価にとどまり、実際の読解行動や読書体験への影響については十分に検証していない。

そこで本研究では、生成 AI を用いて英語教材の可読性を学習者の読解レベルに応じて調整し、多読を支援する手法を提案する。評価実験により、生成 AI を用いた教材書き換えが、学習者の読解レベルに適合した可読性向上を実現できるか(RQ1)、およびそのような可読性調整が読解行動や読書体験にどのような影響を与えるか(RQ2)の検証を目的とする。

2. 生成 AI による可読性向上

提案手法の全体像は図 1 に示すように、本手法は読解力の推定とその結果に基づいた生成 AI による教材の可読性調整から構成される。

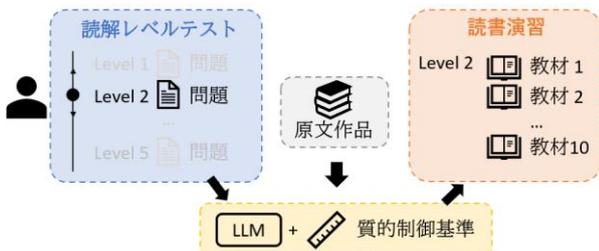


図 1 手法の全体像

2.1 読解力の推定

本手法では、適応型テストの考え方を参考に実装した読解レベルテストにより、学習者の英語読解レベルを推定する。このテストでは、学習者が正解すれば 1 つ上のレベルの問題を、不正解であれば 1 つ下のレベルの問題を提示することで、読解力の推定を効率的かつ負荷が最小限になるように行った。

2.2 生成 AI による可読性向上

次に、推定された読解レベルに基づき、生成 AI を用いて英語教材の可読性を調整する。本研究では Flesch Reading Ease Score⁽⁶⁾を目標値として与え、文章の難易度を数値的に制御する。一方で、可読性指標のみでは語彙の難しさや構文・古風な表現といった質的要因を十分に反映できない。そこで、語彙選択や構文の単純化、古風表現の現代化といった指示をプロンプトに組み込み、主観的にも読みやすい教材の生成を目指した。

2.3 読書演習システムの設計と実装

生成 AI により可読性を調整した教材を用いた読書活動を評価するため、Web ブラウザ上で利用可能な読書演習システムを実装した。教材は約 300 語ごとにページ分割して提示し、ページ移動、読了、リタイアなどの操作を可能とした。読書中の操作はすべて行動ログとして記録し、評価実験に用いた。

3. 評価実験及び分析結果の考察

3.1 実験概要

被験者は英語を第二言語とする大学生 16 名であり、生成 AI により可読性を調整した教材を用いる実験群と、書き換えを行わない原文教材を用いる対照群に無作為に割り当てた。実験に先立ち、読解レ

ベルテスト(15分)を実施し、各被験者の読解レベルを推定した。読書演習(60分)では、両群とも Web ブラウザ上の同一システムを用い、表示形式や操作方法を統一した条件で教材を提示した。

3.2 分析結果と考察

評価指標として、読書演習中の行動ログと、演習後アンケートでの主観評価を用いた。まず、提示された教材が「自分の読解レベルに適しているか」という観点から、文の長さ、語彙の難しさ、構文の複雑さ、古風な表現の有無など、読解難易度に影響を与える要因の主観評価について群間で片側 t 検定を行った。その結果、生成 AI により可読性を調整した教材を用いた実験群は、原文教材を用いた対照群と比較して、有意に高い評価を示した(表1)。この結果は、Flesch Reading Ease Score による数値的な制御に加え、語彙選択や構文の単純化、古風表現の調整といった質的な可読性制御を行ったことが、学習者の主観的な読みやすさの向上に寄与したことを示唆している。

表1 教材に関する主観評価の群間比較

	平均値	標準偏差	最大値	最小値
実験群	3.161	0.400	3.857	2.571
対象群	2.696	0.480	3.286	1.714

($t = 2.10, p = 0.027, \text{Cohen's } d = 1.051$)

※両群ともに N=8

次に、読書体験における没入感について、フロー体験に基づく指標について片側 t 検定を分析した。フロー体験は学習者の技能と活動の難しさの適合によって起こり⁽⁷⁾、読解の継続と習慣化を促すとされている⁽⁸⁾。その結果、実験群は対照群と比較して有意に高いフロースコアを示した(表2)。このことから、生成 AI による可読性調整は、学習者の読解技能と教材難易度との適合を改善し、読書中の没入感を高める可能性があると考えられる。

表2 フロースコアの群間比較

	平均値	標準偏差	最大値	最小値
実験群	4.313	0.826	5.667	3.250
対象群	3.677	0.568	4.583	2.750

($t = 1.79, p = 0.047, \text{Cohen's } d = 0.897$)

※両群ともに N=8

一方、読了ページ数や読書速度、リタイア回数などの行動ログ指標においては、有意な差は確認されず、リタイア回数を除いて実験群の方が低い結果となった(紙面の関係上、詳細な統計データを省略する)。しかし、これは必ずしも否定的な結果を意味するものではなく、実験群においては、より内容理解や没入を重視した慎重な読書が行われていた可能性が考えられる。

また、没入感および読解難易度に関する主観評価

において有意差が得られたことから、本手法は読書量や速度を直接的に増加させるものではなく、読書体験の質や「読み続けやすさ」を支援する効果があるとわかった。

4. まとめ

本研究では、生成 AI を用いて英語教材の可読性を学習者の読解レベルに応じて調整し、多読を支援する手法を提案した。読解レベル推定、生成 AI による可読性制御、および読書演習システムを統合した評価実験を行い、生成 AI による可読性向上が読書体験に与える影響を検証した。

評価実験の結果、生成 AI によって可読性を調整した教材は、原文教材と比較して、読解難易度に影響を与える要因に関する教材の主観評価およびフロー体験に基づく没入感の評価において有意に高い結果を示した。これにより、生成 AI による可読性調整が、学習者の読解レベルと教材難易度との適合を改善し、より没入的で安定した読書体験を促進する可能性が示唆された。

一方で、読了ページ数や読書速度といった行動量指標においては有意な差は確認されなかった。このことから、本手法は多読における読書量の増加を直接的に支援できるものではないが、読書体験の質や継続性を促して多読の継続を支援する手法としての可能性を示した。今後は、被験者数や実験期間を拡大し、長期的な多読実践における学習効果への影響を検証することが課題である。

参考文献

- (1) Julian Bamford, and Richard Day: "Extensive reading: What is it? Why bother?.", *The Language Teacher* 21.5 (1997)
- (2) Richard Day, Julian Bamford, Willy A Renandya, et al.: "Extensive reading in the second language classroom.", *Relc Journal* 29.2, pp.187-191(1998)
- (3) 西澤一, 吉岡貴芳, 伊藤和晃: "英文多読による工学系学生の英語運用能力改善.", *電気学会論文誌 A(基礎・材料・共通部門誌)* 126.7, pp.556-562(2006)
- (4) Chieh-Yang Huang, Jing Wei, Ting-Hao Kenneth Huang: "Generating educational materials with different levels of readability using LLMs.", *Proceedings of the Third Workshop on Intelligent and Interactive Writing Assistants*.(2024)
- (5) Elias Hedlin, Ludwig Estling, Jaqueline Wong, et al.: "Got It! Prompting Readability Using ChatGPT to Enhance Academic Texts for Diverse Learning Needs.", *Proceedings of the 15th International Learning Analytics and Knowledge Conference*.(2025)
- (6) Rudolf Flesch: "A new readability yardstick.", *Journal of Applied Psychology*, 32(3), pp.221-233(1948)
- (7) Kevin Rathunde: "A comparison of Montessori and traditional middle schools: Motivation, quality of experience, and social context.", *Namta Journal*, 28.3, pp.12-53(2003)