

ChatGPT と英語母語話者の英文エッセイの比較

特徴語と前置詞使用の差異に着目して

藤原隆史

1. はじめに

2022 年 11 月に OpenAI 社から ChatGPT が発表されたが、それ以降、世界中で人間社会に大きなインパクトを与え続けている。ChatGPT は、人間が書いた文章と区別がつかないほど自然な文章を生成することができると言われており、言語活動のみならず、言語に纏わる教育活動、すなわち、言語教育にも多大な影響があると考えられている。日本においても、外国語としての英語が習得必須の時代にあつて、ChatGPT によって生成された英文が英語教育に与える影響の大きさは火を見るよりも明らかであろう。本研究では、ChatGPT が生成した英文エッセイを英語母語話者が書いたエッセイと比較し、その言語的特徴を明らかにすることを目的としている。さらに、ChatGPT による英文エッセイが英語ライティングにおける argumentative essay のモデルエッセイとして活用可能かどうかについて考察を行う。

2. 研究の背景

前述した通り、ChatGPT は経済、働き方、言語創作物、教育、研究など、あらゆる分野に影響を及ぼすと考えられており、Eloundou et al. (2023) が指摘するところによれば、米国では労働力の 8 割が少なくとも業務の 10% で何らかの影響を受け、さらに、約 19% の労働者は業務の 50% で何らかの影響を受けると考えられている。また、教育分野でも ChatGPT のような AI の影響を受ける職業の上位に英語を含めた外国語教師が挙げられている（朝日新聞, 2023 年 6 月 11 日記事）。さらに、各大学でも生成 AI の利活用について議論されているが、文部科学省の指針では、「教員による教材開発や、効果的・効率的な大学事務の運営等に利活用することもある」（文部科学省, 2023, p. 2）とされており、考え次第では生成 AI の有効活用を妨げるものではないと考えられる。

以上の状況を踏まえ、本研究では ChatGPT を英語教育において有効活用することを視野に入れ、ChatGPT が生成した英文エッセイを英語母語話者が書いたエッセイと比較し、ChatGPT によるエッセイの言語的特徴を明らかにすることを目的とする。具体的には、同一のエッセイ課題（レストランでの喫煙の是非）について、ChatGPT と英語母語話者のエッセイを比較する。その上で、ChatGPT が生成したエッセイが英語試験などでしばしば見られる argumentative essay のモデルエッセイとして活用可能かどうかについて検証する。

3. 研究方法

本研究では、GPT3.5 ベースの ChatGPT（無料版）を用いて分析を行った。現在 GPT4.0 をベースとした有料版が公開されているが、本研究の目的である英文の生成は、GPT3.5 ベースの ChatGPT でも問題なく行うことができると判断し、有料版の ChatGPT は用いていない。比較の対象となる英語母語話者によるエッセイは、神戸大学の石川慎一郎氏による The International Corpus Network of Asian Learners of English (ICNALE) に収録されているものを用いた。また、エッセイ課題としては「レストランでの喫煙の是非」をテーマとして書かせる課題で、指示文章（prompt）は Ishikawa (2013) による指示文章を一部改変した以下の prompt を用いた。

Do you agree or disagree with the following statement? Use reasons and specific details to support your opinion.
Statement: Smoking should be completely banned at all the restaurants in the country. When writing your essay, clarify your opinions and show the reasons and some examples. The length of your essay must be from 200 to 300 words.
(Ishikawa, 2013: 一部改)

分析の対象は、reference corpus として ICNALE に収録された英語母語話者によるエッセイ 200 篇 (ENS_SMK) (44839 words)、target corpus として ChatGPT が生成したエッセイ 30 篇 (GPT_SMK) (19445 words) を設定した。さらに、分析のためのツールとしては、コンコーダンスとして AntConc 4.2.0、語彙レベル判定に New Word Level Checker (NWLC) (Word List を New JACET8000 に設定)、語彙数集計に Voyant Tools を用いた。

4. 結果と考察

先ず、コーパス言語学の分析方法に従い、各コーパスにおける平均 token 数、平均 type 数、Type /Token Ratio (TTR)、1 文の平均語数を示す。GPT_SMK の平均 token 数は 388.9、平均 type 数は 190.24、TTR は 49%、1 文の平均語数は 20.9 語であった。ENS_SMK の平均 token 数は 224.1、平均 type 数は 126.9、TTR は 57%、1 文の平均語数は 22.8 語であった。主なポイントとしては、ChatGPT は文字数の指定を無視し、平均 token 数と type 数が多いという結果となった。TTR は英語母語話者の方が高いが、コーパスサイズの違いを考慮して TTR の修正指標である R を求めたところ、GPT_SMK が $R=9.6$ 、ENS_SMK が $R=8.4$ となり、GPT_SMK の語彙密度の方が高くなった。また、NWLC で語彙レベルを調べたところ、GPT_SMK の方が学習者向けに作られた New JACET8000 のリストにおける語彙レベルが高い (Level 8 の段階での語彙蓄積率が、GPT_SMK で 97.73%、ENS_SMK で 98.31%) 結果となった。語彙密度との関連でいえば、GPT_SMK は New JACET8000 に搭載されている語彙ではレベルが高めで語彙密度も高いため、学習者にとっては ChatGPT が生成した英文の方が読解しにくいと判断できる。しかし、語彙リストに搭載されていない語 (Not Available words) を見てみると、GPT_SMK の NA words が 79 語であるのに対し、ENS_SMK の NA words は 377 語であった。すなわち、語彙リストにないような極端に難しい語や低頻度語も考慮に入れると、ENS_SMK の読解難易度の方が高い可能性がある。換言すれば、ENS_SMK は不均一でばらつきのある文章となっている可能性があるが、GPT_SMK は語彙リストにある程度搭載された語が用いられたバランスの取れた文章であり、読解難易度が相対的に低くある程度統制された語彙学習を想定した英文を生成させられる可能性が示唆された。

次に、GPT_SMK の特徴語と特徴的低頻度語について述べる。ChatGPT が生成したエッセイは、最初に「レストランでの喫煙の是非」について賛成か反対かを表明し、introduction-body-conclusion という形式でエッセイを書いており、argumentative essay として書かれていることが分かる。(そのため、特徴語上位には、conclusion, paragraph, body, title がきている。)しかし、GPT_SMK の 30 篇のエッセイ全てがレストランでの喫煙に「反対」の立場で論を展開しており、生成された内容に偏りが生じている。そのため、特徴語としては secondhand smoke, ban, respiratory problem/ailment, prioritize public health, create a healthier environment, health risks, promoting a smoke-free environment/ a healthier dining experience, ensure a safe and pleasant dining environment, protect the health of employees/patrons といったものがきており、喫煙に賛成するエッセイを生成させたい場合は prompt を工夫する必要がある。また、前置詞 by の使用に特徴があることも分かった。すなわち、by implementing a complete/nationwide ban, by eliminating (secondhand) smoking, by banning smoking (in restaurants), by prohibiting smoking, by enforcing a smoking ban, by eliminating the health hazards/ smoking, by creating/establishing smoke-free environments, by going smoke-free, by protecting public health 等のフレーズが多用されており、「喫煙が有害であるから〇〇をすることで解決すべき」という plot で by doing... という副詞句が多く用いられたことが分かった。この点について、モデルエッセイとして学習者に示す際は、当該表現の定着には有用であると考えられるが、その一方で前置詞使用に偏りがあり、包括的な学習という観点からは疑問が残ると言わざるを得ない。

さらに、特徴的低頻度語については、人称代名詞、do, be, have、if 節、否定語、feel, think, believe 等の主観的表現、といった、自己に言及する表現を避け、客観的表現を用いる傾向があることが分かった。すなわち、客観性が高く argumentative essay のモデルエッセイとして有効活用できる可能性があることが示唆された。

5. 結論

本研究では、ChatGPT が生成した英文エッセイを英語母語話者のエッセイと比較し、argumentative essay として活用可能かを検証した。結果として、prompt を工夫したり生成された英文を教師がチェックしたりすることでモデルエッセイとして活用できる可能性があることが示された。

参考文献 (一部)

「英語教育、チャット GPT で変わる? 誰もが英語“使える” ツール登場——専門家に聞く」『朝日新聞』(2023 年 6 月 11 日)/Eloundou T, Manning S, Mishkin P, Rock D (2023) “Gpts are gpts: An early look at the labor market impact potential of large language models”, arXiv preprint arXiv 2303.10130./ Ishikawa SI (2013) “The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English”, *Learner corpus studies in Asia and the world* 1(1), pp. 91–118./ 文部科学省 (2023) 「大学・高専における生成 AI の教学面の取扱いについて (周知)」, https://www.mext.go.jp/content/20230714-mxt_senmon01-000030762_1.pdf (2023 年 8 月 14 日閲覧)/ New Word Level Checker, <https://nwlc.pythonanywhere.com/> Voyant tools, <https://voyant-tools.org/>.