

AIのある英語教育・研究 Let There Be AI!

水本 篤 関西大学

1. 生成AIの広まり

生成AIとは、人工知能(AI)が人間のように各種コンテンツを作り出す技術を指す。AIは写真、動画、音声、音楽、文章、プログラムコードなどを生成することが可能である。2023年は生成AI元年であったと言っても過言ではない。OpenAIが2022年11月30日にChatGPTを公開して以来、わずか2カ月で1億人のアクティブユーザーを獲得し、世界中で大きな反響を呼んだ。そのため、2023年に入ってから、国内外の大学をはじめとして、多くの教育機関、そして文部科学省が生成AIの利用に関するガイドラインを発表した。

生成AI、特にChatGPTは、さまざまな研究分野にも影響を及ぼしている。応用言語学分野で2023年前半に出版された論文では、ChatGPTを言語教育と学習に使用していく可能性や(Kohnke et al., 2023)、ChatGPT出現後の学術出版における課題や問題点を指摘しているもの(Teng, 2023)が多く、実証研究は一部(e.g., Mizumoto & Eguchi, 2023)を除いては少なかった。現在は、ChatGPTが教育現場で受け入れられつつある状況を反映し、学習者がChatGPTを使用した結果を報告する実証論文(e.g., Liu & Ma, 2023; Yan, 2023)も増えている。著者が査読している海外ジャーナルに投稿されるChatGPTをもちいた論文も、毎月4~5本査読しなければならないペースで増えており、2024年以降は、さらにChatGPTなどの生成AIを教育や学習で使用し、その効果を報告する論文がさらに増えるだろう。

文章の生成を行うことができる生成AIは、大規模言語モデル(Large Language Model: LLM)をもちいて、さまざまな質問に対する回答を行うチャットボット形式でAIとのやりとりを行う。LLMは語法・文法的に正確な文を生成することができ、文法説明

や分析方法、コードの説明も可能であるため、学習者が利用して学習効果を高めるだけでなく、教員が教材開発や研究活動を効率化するためにも使用できる。LLMには(Chat)GPTだけではなく、Gemini(旧Bard)、Claude、LlamAなどの競合する言語モデルも存在し、これらはそれぞれ独自の特徴と応用領域を持つ。AI技術の進化により、これらのモデルは今後、さらに多様な用途に対応する能力を持つようになるだろう。本稿では、2024年2月現在、最も性能が高いとされる有料版のGPT-4を搭載したChatGPTを例に挙げて説明を行う。

2. GPTの仕組み

ChatGPTはLLMをもちいている。LLMは、大量のテキストデータ(ウェブページ、書籍、記事など)から意味や文脈理解を含めた形で言葉の使用方法を事前に学習し、その特徴をベクトル(vector)として数値化として格納している。学習データは、それぞれの単語がどのように文脈内で使用され、次の単語とどのように関連するかを理解するように設定されており、この事前学習には「深層ニューラルネットワーク」(Deep Neural Network: DNN)がもちいられている。このDNNは「Transformer」と呼ばれる技術を利用しており、Google翻訳やDeepLなどの高品質な機械翻訳にも使われている。事前学習を基に、ユーザーからの文字入力に対して、学習した大量のデータを基にして続きの文章を確率的に予測し生成したり、質問に答える能力を持つ。指示文の理解、適切な表現の使用、不適切な発言の抑制などを組み合わせることで、自然な対話が可能になっている。文章は人間を超える流暢さで、長文でもほとんど破綻がなく、さまざまな指示に対応できる。このようなLLMの一つがGPT(Generative Pre-trained

Transformer)である。

自然言語処理(Natural Language Processing: NLP)の分野で使われる「分散表現」(word embedding)は、単語や文章などの言語的要素を高次元の連続ベクトル空間にマッピングすることを指す。この分散表現により、言語的要素の意味的、文法的特徴などがベクトルにエンコード(符号化)される。この技術により、単語を関連性や類似度に基づいてベクトル化し表現することが可能になる。たとえば、word2vec (Mikolov et al., 2013)のようなモデルは10年以上前から存在するが、大きなブレークスルーはTransformer (Vaswani et al., 2017)の登場によるものである。特にGoogleのBERT (Bidirectional Encoder Representations from Transformers) (Devlin et al., 2018)や、OpenAIのGPT (Radford et al., 2018)のような高度な文脈依存型の言語モデルの出現により、単語の周囲の文脈を考慮した精密な分散表現を生成することができるようになり、今日のChatGPTの形につながった。

Transformerの仕組みをわかりやすく可視化して説明している例がFinancial Timesで紹介されており(<https://ig.ft.com/generative-ai/>)、その中からいくつかの例を以下に示す。Transformerのself-attention(自己注意)によって、LLMは文中の単語間の関係を前から一つずつではなく、一度に全体として理解することができる。たとえば、(1)では、前の「no」と後ろの「in」に注意を向けることで、「interest」が「興味」という意味の名詞であることを理解できる。この仕組みがTransformerで最も重要なself-attentionの特徴である。

(1) I have no interest in politics.

文を長くして、(2)のようにしてみても、全体と前後の文脈が処理され、2つ目の「interest」が「金利」を意味していることが、「rate」や「bank」(あるいは前の「rising」)に注意を向けることで理解できる。

(2) I have no interest in hearing about the

rising interest rate of the bank.

このようなself-attentionの仕組みによって、同義語や反意語、また、上記の「interest」の例のように多義語でも文脈に沿った理解や生成が可能になる。文脈全体を理解できる仕組みは、以下の(3)や(4)のような例において、文脈に応じて代名詞の参照先を変えることができることから明らかである。

(3) The dog chewed the bone because it was hungry.

(4) The dog chewed the bone because it was delicious.

self-attentionは、単一の文だけでなく、段落やパッセージ、さらには本全体の情報を保持しながら単語の意味を理解する能力を持つため、学習に使われるデータ(コーパス)の規模が大きくなるほど、その精度が高まる。人間が文を理解するのと同じような仕組みを、LLMではTransformerのself-attentionメカニズムによって効率的に実装しており、これまでにないレベルの精度と理解度を実現している。たとえば、機械翻訳においてTransformerモデルは、文脈を深く理解し、より自然で流暢な翻訳を可能にしている。従来の翻訳モデルが単語やフレーズレベルの変換に留まっていたのに対し、Transformerモデルはより広い文脈を考慮し、言語間の微妙なニュアンスや文化的な差異を捉えることができる。2017年に発表されたTransformerの仕組みを紹介した論文のタイトルは、“Attention is All You Need” (Vaswani et al., 2017)であり、実際にattentionがLLMの核となっていることがわかる。Transformerモデルのアプローチが、その性能の高さから、さまざまな応用分野で今後も広まっていくと思われる。

3. ChatGPTなどのLLMが得意なこと・不得意なこと

前項で説明した通り、LLMの基本的な仕組みは、事前に学習した大量のデータを基に、続きの単語や

文章を確率的に予測し生成することである。これは、単語の位置情報や文脈の情報を全て学習した言語モデルであり、加えて自らが学習したパターンから新しい文章を生成する能力も持つ。「文理解と生成ができる辞書」と考えると理解しやすいかもしれない。そのため、言語に関するタスクでは非常に優れており、研究や実践でアシスタントとして使うことができる。例えば、手作業で行われるコーディングやアノテーション(メタデータのタグ付け)のサポートも可能である。実際には、クラウドソーシングで人間がアノテーションを行った場合と比べて、ChatGPTが行ったアノテーションの方が正確であるという報告もある(Gilardi et al., 2023)。

一方で、確率的に次の語を生成する仕組みを考えると、数を数えたり、計算することは苦手である。特に、語数のカウントにおいては、Microsoft Wordのように英文の語数をカウントさせると正確な数値が表示されることはほとんどない。そのため、CAF(Skehan, 1996)として知られる、複雑さ(complexity)、正確さ(accuracy)、流暢さ(fluency)といった指標を計算させることはできない。文中のT-unitや文法・語法の誤りを指摘し、説明することはできても、合計頻度の計算ができないからである。この事実を知らなければ、Lin(2023)が誤って提案しているように、「頻度表をChatGPTが作成できる」と報告してしまう可能性がある。

さらに、ChatGPTは言語モデルであり、正確な・最新の情報を検索する目的で作られていないため、Google検索のように使うことはできない。ChatGPT有料版のGPT-4やMicrosoftのCopilot(旧Bing Chat)はWebから最新情報を回答に含めることができるが、ハルシネーション(幻覚)を起こして虚偽の情報を提供する可能性もある。そのため、回答の検証が重要である。論文の検索や発見が目的であれば、Perplexity AI(<https://www.perplexity.ai/>)のような特化したサービスを使用する方が良いだろう。

また、ChatGPTは平均的で無難な文生成を行うよう人間のフィードバックに基づいてトレーニングさ

れているため、専門知識が必要なタスクには向いていない。何らかのコーディングを行う際、ChatGPTが一から行うと凡庸なカテゴリーの提案に留まることがあるが、ある程度のカテゴリーが決まっている場合には、それに合わせたコーディングを正確に行うことができる。つまり、人間が作成したカテゴリーに対して、対応するテキストデータを抽出する能力は優れている。

これらの特性から、「ChatGPTは実践や研究では使えない」と結論付けるのは早計である。ChatGPTを辞書やGoogle検索、機械翻訳ツールの延長線上のツールと考え、その利点を活用すれば、非常に有用なパートナーになるだろう。

4. 英語教育実践でのChatGPTの使用

ChatGPTは幅広い言語理解・生成タスクが得意であり、教室内外での実践においてさまざまな用途が考えられる。以下に挙げるような、「これはChatGPTにやらせてみたらできるのでは?」というほとんどのことが可能である(水本, 2023; 田村, 2022)。

- ・ 翻訳
- ・ 英文(教科書本文)の加工、要約、パラフレーズ
- ・ テストなどでの(たたき台としての)英文の自動作成
- ・ ライティング(あるいはスピーキング原稿の)サンプルの提示
- ・ ライティングの採点とフィードバック
- ・ 教室内的パフォーマンス・テストのルーブリック作成補助
- ・ 教室外での学習のサポート

著者自身、「教科書本文の新出語を使って別の文章を作る」という課題を、コーパスやGoogle検索など多様な方法で試みてきたが、ChatGPTの登場によってこの課題が一気に解決した。卯城(2012)が提案しているように、教科書本文のダイアログをモノログに変えたり、類語・類似表現に書き換えたり、「付け足し」や「まとめ」を加えるなど、教科書の内容をパラフレーズすることがChatGPTを利用するこ

とで可能になる。これまでは英文チェックをALTに依頼していたが、ChatGPTの登場によりその必要もなくなった。さらに、ChatGPTはAIであり、人間のような複雑な(そしてややこしい)感情は持っていないため、何度も質問しても申し訳ないと感じる必要がない。

英語教育実践においてChatGPTを最も効果的に活用できるのは、ライティングの採点とフィードバックである。Mizumoto and Eguchi (2023)の研究では、TOEFL受験者12,100名のエッセイをGPT (text-davinci-003)で自動採点し、その結果が人間の採点とある程度一致していた。単独での使用では精度がそれほど高くないものの、人間の採点のサポートとして利用することは可能であると報告されている。実際には、ChatGPTを使って学習者が書いたものをCEFRのレベルを指定しながら自動採点させたり、文法や構文、語法の誤りに対するフィードバックも出力させることが可能である。また、学習者が書いた間違いを含む英文を正確に書き直してもらうことで、教員が添削する必要がなくなり、学習者は自らの誤りに対するフィードバックを即座に受け取ることができる。これにより、従来とは異なる効果的なライティング指導が可能になる。ChatGPTを使ってライティングのフィードバックを行う際には、表形式で提示したり、日本語での説明をChatGPTに書いてもらうことも可能であり、教えている学習者のレベルに応じて適切なフィードバック方法を教師が考えることになる。

今後は、ChatGPTを含む生成AIがテキスト、画像、音声、動画などの複数の種類の情報を一度に処理できるマルチモーダルAIに進化していくことを考慮に入れると、近い将来、教室内外でChatGPTなどを用いてスピーキングの即興性を高める練習も可能になるだろう。自動会話システムであるチャットボットの即時性、使いやすさ、個人最適化、不安軽減の観点からその効果が報告されている(Huang et al., 2022)。これらの観点から、学習者のスピーキング能力を向上させるサポートツールとしても、ChatGPTの効果的な利用が期待される。

5. プロンプトの重要性

ChatGPTが世界中で大人気になり、広まった大きな要因の一つが、プロンプト(prompt)によって指示を出せることである。これにより、従来のようにコードを書いてプログラムに指示を出すことなく、人間が使う自然言語によってChatGPTと直接会話形式でやり取りができるようになった。ChatGPTのようなLLMでは、プロンプトをうまく設計することがタスクを遂行するために重要である。適切な指示を入力し、最適な出力を得るための技術を「プロンプトエンジニアリング」(prompt engineering)と呼び、この技術には、意図した出力を得るために、質問をどのように構築し、どの情報を含め、どのように情報を提示するかが含まれる。例えば、具体的な事例や状況の説明、特定の言語スタイルやトーンの使用、そして必要な情報の適切な指定などが重要である。

会話形式でやり取りができることは、一般ユーザーにとってアクセスしやすくなるが、思ったような回答を得るためには、プロンプトの工夫が必要である。また、ChatGPTは内部の設定で毎回異なる回答を生成するため、プロンプトの作成は、コンピュータ・プログラムのコードを書く際のトライアル・アンド・エラーと同様、重要となる。同時に、ChatGPTが理解しやすいように指示を出すため、ユーザー側にもコミュニケーション能力が求められる。一般的には、プロンプトを設計する際には、詳細な説明を与え、明確かつ正確な指示を出し、いくつかの具体例を含めることによってパフォーマンスが向上することが知られている(Brown et al., 2020)。Bsharat et al. (2023)は、これらの原則を整理・検証し、プロンプト作成の原則26項目をまとめており、とても参考になる。

英語教育実践と研究において、自分の実践や研究に適したプロンプトを蓄積することは極めて有益である。報告を行う際には、内容だけでなく、使用したLLMの種類とプロンプトも併せて提示することが肝要である。優れた実践に加えて、プロンプトの成功例も広く共有することが推奨される。

例えば、柳瀬陽介先生（京都大学）が作成し公開している「ChatGPT 学術英語ライティング添削・改訂プロンプト」(<https://yanase-yosuke.blogspot.com/2023/11/ver4chatgpt-3.html>)、『ChatGPT × 教室の仕事』（2023年，明治図書）の著者、南部久貴先生が公開している「英語教師のための ChatGPT プロンプト集」(<https://github.com/nmb1019/chatgpt-for-english-teachers/>)、高木俊輔先生（聖光学院中学校高等学校）の「英語教員のための ChatGPT プロンプト集」(<https://bit.ly/prompts-for-english-teachers>)などは貴重なリソースとして大いに役立つ。

また、本稿の著者も、ユーザー自身がプロンプトを保存し、他者と共有できるサイト (<https://langtest.jp/me-mo/>) を構築し、無償で公開している。このサイトを一度訪れ、英語教育実践・研究において活用可能なプロンプトがあれば、ぜひその情報を登録し、他のユーザーと共有していただきたい。

6. 今後の展望

生成 AI 元年であった 2023 年を振り返ると、日進月歩の生成 AI への期待と希望に満ちた 1 年だったと感じられる。しかし同時に、「ChatGPT があれば英語教員は不要になるのでは？」という悲観的な意見も耳にした。本稿をお読みの方ならお分かりいただけると思うが、ChatGPT のような生成 AI は人間の指示が必要であり、まだそのような段階には至っていない。ただし、ChatGPT に取って代わられるような、言語知識のみに焦点を当てた授業を行っている場合は、教授法を根本から変える必要がある時期にきていることは間違いない。

これからの時代を生き抜く学習者を育てるためには、「生成 AI を使用しない」という選択肢は現実的でないと言える。学習者も AI を活用する社会で生活していくことになるため、AI との共存は不可欠である。教育者としての「愛情」があれば、教室に「AI」を導入することを当然と考え、積極的に活用していくべきである。そのためにも、私たち英語教員は生成 AI を積極的に取り入れ、教室内外での活用法を模

索するべきである。また、生成 AI を効果的に使用するためには、出力が正しいかどうかを判断する能力が求められるため、英語運用能力が必要になる。つまり、英語教師として英語自体を学び続け、研鑽を積むことが今後さらに重要となる。AI 全盛の時代に「英語教員は AI があっても絶対に必要である」と自信を持って言えるようになるために、自己研鑽を続けよう。（自戒の念を込めて。）

引用文献

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *arXiv*. 2005.14165. <https://doi.org/10.48550/arXiv.2005.14165>
- Bsharat, S. M., Myrzakhan, A., & Shen, Z. (2023). Principled instructions are all you need for questioning LLaMA-1/2, GPT-3.5/4. *arXiv*, 2312.16171. <https://doi.org/10.48550/arXiv.2312.16171>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv*, 1810.04805. <https://doi.org/10.48550/arXiv.1810.04805>
- Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd-workers for text-annotation tasks. *arXiv*. 2303.15056. <http://arxiv.org/abs/2303.15056>
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237-257. <https://doi.org/10.1111/jcal.12610>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023).

- ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>
- Lin, P. (2023). ChatGPT: Friend or foe (to corpus linguists)? *Applied Corpus Linguistics*, 3(3), 100065. <https://doi.org/10.1016/j.acorp.2023.100065>
- Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*. <https://doi.org/10.1080/17501229.2023.2240316>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *arXiv*, 1310.4546. <https://doi.org/10.48550/arXiv.1310.4546>
- 水本 篤 (2023, May 27). 「そこに AI はあるのか? - ChatGPT による外国語教育実践・研究の革新」外国語教育メディア学会 (LET) 関西支部 2023 年度春季研究大会ワークショップ資料 <https://bit.ly/3ME7cPz>
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050. <https://doi.org/10.1016/j.rmal.2023.100050>
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. *OpenAI Preprint*. <https://arxiv.org/abs/1805.06905>
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17(1), 38-62. <https://doi.org/10.1093/applin/17.1.38>
- 田村 祐 (2022, December 13). 「ChatGPT で英語授業の教材づくりをしてみる」[Blog post]. Retrieved from <https://tam07pb915.com/2022/12/13/chat-gpt-and-teaching-materials/>
- Teng, F. (2023). Scientific writing, reviewing, and editing for open-access TESOL journals: The role of ChatGPT. *International Journal of TESOL Studies*, 5(1), 87-91. <https://doi.org/10.58304/ijts.20230107>
- 卯城 祐司 (編著) (2012). 「英語リーディングテストの考え方と作り方: 教科書とテストをつなぐ指導と評価」研究社
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is all you need*. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. <https://doi.org/10.48550/arXiv.1706.03762>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28, 13943-13967. <https://doi.org/10.1007/s10639-023-11742-4>



水本 篤(みずもと あつし)

関西大学 外国語学部・外国語教育学研究科 教授。専門はコーパスの教育利用、語彙学習方略、言語テスト。近年はコーパスを利用した学習と指導に関する研究を主に進めている。英語コーパス学会誌『英語コーパス研究』編集委員長。Applied Corpus Linguistics 編集委員。著書に、『外国語教育研究ハンドブック【増補版】— 研究手法のより良い理解のために』(2023年、松柏社、共編著)、『ICTを活用した英語アカデミック・ライティング指導—支援ツールの開発と実践—』(2017年、金星堂、編著)などがある。2014年全国英語教育学会(JASELE)学術奨励賞受賞。2016年英語コーパス学会(JAECS)奨励賞受賞。2017年外国語教育メディア学会(LET)学術賞受賞。趣味はPython(永遠の日曜プログラマー)と黒ラブのメイちゃんを愛でること。