

ライティングツールを利用した英作文添削比較

甲斐 順

神奈川県立松陽高等学校 〒245-0016 神奈川県横浜市泉区和泉町 7713

E-mail: vyg03722@nifty.com

あらまし 本研究の目的は、生成 AI を活用したライティングツールを複数用いて、高校生が書いた英作文の添削を比較、分析、考察し、英作文指導への教育的示唆を行うことである。57 人の英作文データを 4 つのライティングツールに分析させ、その結果を検証した結果、ツール数が増えるにつれて修正案の一致数は減る傾向にあり、一致する修正項目は減っていた。ライティングツールが修正案を提示せずに修正を行う実態や未修正のままの部分が見られる実態も判明した。

キーワード 英作文, ChatGPT, DeepL Write, English Level Checker, Grammarly

A Comparison of the Correction of English Writing with the Help of Writing Tools

Jun Kai

Kanagawa Prefectural Shoyo Senior High School 7713 Izumi-cho, Izumi-ku, Yokohama-shi, Kanagawa, 245-0016

Japan

E-mail: vyg03722@nifty.com

Abstract The present study aims to analyze and compare what senior high school students wrote in English, using various writing tools. It seeks to assess the outcomes and provide pedagogical implications. Data were collected from fifty-seven students and analyzed. The findings reveal that as the number of writing tools increased, there was a tendency for a decrease in the consistency of error suggestions among these tools. Additionally, the results indicate that certain writing tools correct errors without providing guidance on the correction process and that error corrections suggested by these tools are sometimes overlooked or disregarded.

Keywords English writing, ChatGPT, DeepL Write, English Level Checker, Grammarly

1. はじめに

学習者の英作文の添削は多大な労力を要する。判読しにくい文字に始まり、綴りの誤り、使用語彙や語法・文法の誤り、意味不明な英文、日本語の発想で結論を後回しにして書かれた英文などに遭遇するたびに、読みが中断され、朱筆をどう入れて推敲し、適切な助言をどのように与えればよいか等、枚挙に暇がない。柳瀬 (2023b) は、英語ライティング指導について、学生の英作文の添削、改定案、ポイントの説明、励ましの言葉まで添えると学生 1 人当たり最低 30 分はかかり、継続して実施すると過労で健康を損ねかねないと述べる。

そんな中、近年 ChatGPT を始めとする生成 AI (Artificial Intelligence) が登場し、日常生活だけでなく、日本の英語教育をも大きく変えようとしている。柳瀬 (2023b) は ChatGPT を使えば、数十人の添削、

フィードバックが短時間で終わり、教員は千人力を得るとまで述べている。教員の利用だけでなく、学習者がライティングツールを用いて英作文を自力で添削することにより、教員の負担が軽減される可能性がある。その前提として学習者がライティングツールの使い方に慣れており、教員にライティングツールの添削について指導できるだけの知見を有することが必要となる。

本研究では、生成 AI を活用したライティングツールを複数用いて、高校生が書いた英作文の添削を比較、分析、考察することで、今後の英作文指導への教育的示唆を行う。

2. 文献レビュー

2.1 ライティングツールに関する先行研究

Barrot (2023) は、Grammarly を用いた自動ライティング訂正フィードバックの効果について、フィリピン

の大学生を対象に検証した。結果として、ライティングの正確性を改善することが判明した。Barrot は、正確性を高める要因として、学習者の誤りに対して Grammarly から提供される様々な修正案への学習者の気づき、順応性のあるメタ言語的な説明、学習者の自律的学習を促すことを挙げている。ただし、学習者の中には、Grammarly の限られた説明や過剰訂正を疑問視する者がいれば、訂正の多さから認知的過負荷を感じる者がいることが報告されている。

Grammarly を用いた Barrot の研究はフィリピン人を対象としているが、日本人英語学習者を対象とした研究も見られる。新美・梅木 (2023) は、日本人大学生に Grammarly を用いて英語ライティングの修正活動を行わせた結果、文法や表記の面で明らかな伸長が見られたと述べている。新美・梅木は、文法や表記に対する意識が高まった要因として、自己分析力が促進された可能性を指摘する。その一方で、学生が Grammarly を活用した修正活動を行っている際に、どのような修正項目をどのように修正しているのかについて検討する必要性も挙げている。

Schmidt-Fajlik (2023) は、日本人大学生が書いた短い英文テキストを ChatGPT、Grammarly、ProWritingAid の 3 つを利用して分析を行った。日本人英語学習者の文法を手助けする点に関して、ChatGPT が最も有益で最も良い解答を提示すること、大学生に課したアンケート結果から、ChatGPT は利用しやすく、綴り、文法など、段落を点検する際に役に立つと実感していることが明らかにされている。ただし、自由回答形式のアンケートの中には、ChatGPT を初めて利用し、少し難しいと実感している学生が複数名いることも示され、どういった点で難しいと感じていたのかまでは明らかにされていない。

大学生以外の日本人英語学習者を対象とした研究に Gold (2023) がある。30 代の日本人女性 1 名を対象に Grammarly と DeepL を用いたライティングの効果を検証したところ、これらのツールは重要なアフォーダンスを提供し、学習の足場掛けとなる支援を行い、認知的負担や不安を低下させ、ライティングに対する自信や動機を高めるということが判明した。興味深いのは、実験参加者が Grammarly の提案には時に欠陥があり、不適切であった、また DeepL の翻訳についても時々不正確であったと報告している点である。Gold (2023) は、Grammarly の提案はいつも正確であるとは限らないので、学習者による批判的な熟慮が必要であり、DeepL についても学習者の批判的思考力やメタ言語的な熟慮が必要であると述べている。

実証研究ではないが、柳瀬 (2023a) は日頃英語で書く経験が乏しい日本語話者向けに、AI を活用した英語

論文を執筆する方法について解説している。DeepL などの機械翻訳 AI が出力する英語は、スペリングや文法へのミスが事実上皆無で、ChatGPT などの多目的の生成系 AI は英語の文体改善もすると述べる。ただし、現時点の AI は、言語表現について万能ではなく、AI が生み出した英文を批判的に精読できることが必要であると主張する。

AI を中心としたライティングツールは学習者にとって有益であることが文献レビューからうかがえるが、必ずしも正確な、あるいは適切な提案を行うとは限らないことが指摘されている。提案を受け入れるかどうかは英文を作成した人間が批判的に熟慮し、決定を下す必要があることがわかる。ただし、習熟度の低い学習者は、提案を鵜呑みにして熟慮せずに受け入れ、提案を採用してしまう可能性があり、適切な支援が必要であるとも言える。

2.2 第二言語ライティングにおける訂正フィードバック

第二言語ライティング研究における訂正フィードバックは、Truscott (1996) が「ライティング授業に文法の訂正の場はなく、廃止すべきである」との主張に端を発し、一躍脚光を浴びるようになった。Ferris (1999) は、Truscott が取り上げた「誤りの訂正」という用語の定義や先行研究に問題点があり、彼の主張は時期尚早で、明らかに強い物言いであるとして、「誤りの訂正」の継続を訴えた。

その後、Bitchener (2008)、Bitchener, Young and Cameron (2005)、Ellis, Sheen, Murakami and Takashima (2008)、Sheen (2007) などのライティングにおける訂正フィードバックの有効性を示す研究が見られるようになる。Hyland, Nicolás-Conesa, and Cerezo (2018) は、こういった一連の流れを受けて、ライティングにおける訂正フィードバックは、学習者のライティングの正確性を向上させるとともにその効果が長期間持続するという立場を支持することが近年の研究結果から明らかになってきていると述べている。

なお Truscott (1996, 1999) も Ferris (1999, 2002) も「誤りの訂正」は学習者が自己編集できるようになるのに有益であるという点で一致している。Ferris (1999) は、訂正フィードバックやストラテジーの訓練がなければ学習者は編集する技量を改善する必要性を真剣にとらえないとさえ述べる。

3. 調査方法

3.1 目的

文献レビューから、生成 AI を活用したライティングツールが示す提案には、綴りや文法など様々あり、それらが適切な提案である場合もあれば不適切な場合

もあることが指摘されている。学習者がライティングツールの提案を批判的に熟慮し、賢明に判断できるようになり、最終的に誤りのない作文を自律的に書けるようになることが望まれる。また、第二言語ライティングにおける訂正フィードバックの文献レビューから訂正フィードバックの有効性が明らかになると同時に学習者が自己編集できるようになるのに役立つことが指摘されている。生成 AI を活用したライティングツールを通して、訂正フィードバックによる正確性の向上や学習者の自己編集力が育まれる可能性がある。1つのライティングツールを使ってこれらが達成される可能性もあれば、複数のライティングツールを使って達成される可能性もある。ただし現時点の AI は、言語表現に関して万能ではない（柳瀬, 2023a）。複数のライティングツールを用いて学習者の英作文を添削させ、提案を含めた修正結果の一致点、一致しない点を比較、分析、考察することにより、学習者が批判的に熟慮し、賢明に判断できるようになり、自律的に正しい作文を書けるようになると思われる。その前提として、指導する教員の側に複数のライティングツールの添削結果を比較、分析、考察し、学習者へ適切な指導を行えることが必要となる。文部科学省（2023）は、「教師が生成 AI が生成する誤りを含む回答を教材として使用し、その性質やメリット・デメリット等について学ばせたり、個人情報や機械学習させない設定を教えることも考えられる」（p. 8）と述べており、教員が生成 AI を活用したライティングツールについての知見がなければ指導することはできない。

そこで本研究では、高校生が書いた英作文を複数の AI を活用したライティングツールに添削させ、ライティングツールが示す提案を含む修正結果の一致点、一致しない点を比較、分析、考察し、今後の英作文指導に資することを目的として行う。次の3つの研究課題を設定する。

研究課題 1 複数のライティングツールが示す修正案はどの程度一致が見られるか。

研究課題 2 複数のライティングツールが示す修正案にはどのような内容が見られるか。

研究課題 3 複数のライティングツールの中には、修正案を示さずに修正している部分は見られるか。

3.2 分析用ライティングツール

本研究では、先行研究で挙げられていた ChatGPT、Grammarly に加えて DeepL Write, English Level Checker を分析用のライティングツールとして用いる。

ChatGPT は、AI を搭載した自然言語処理プラットフォームで、人間のような反応を生成するために機械学習と自然言語処理の組み合わせを利用している

(Schmidt-Fajlik, 2023)。ChatGPT は、様々な言語に関連するタスクで非常に優れており、機械翻訳、英文の加工、要約、パラフレイズ、ライティングの採点とフィードバックなど様々な用途が考えられる（水本, 2023）。ライティングの採点とフィードバックでは、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の自動採点、文法や構文の誤りの指摘、適切な語彙の提案などのフィードバックを出力させることも可能である（水本, 2023）。

Grammarly は、作成された文章について、文法や表記、文章の一貫性や明瞭さ等を添削し、誤りや不適切な表現に色つきの下線が引かれ、修正案を提示する（新美・梅木, 2023）。

DeepL Write は AI 搭載の文章作成アシストツールで、文法や句読点、文体をチェックし、提案や書き換え候補を提示して、文章作成をアシスト、英語（アメリカ/イギリス）、ドイツ語に対応し、左側の入力欄にテキストを入力すると、右側の出力欄に推敲済みのテキストを表示、出力欄に緑色で表示された部分が変更箇所を示し、書き換え候補を参照して、単語や文全体を書き換えることが可能である（DeepL, 2022）。

English Level Checker は CEFR のレベルを付与したテキストの語彙・文法を中心とした特性を AI が学習し、教科書・教材のテキストレベルの判定を行い、学習者の英作文のレベルと誤りを判定することができる（投野, 2023）。英作文をチェックしたい場合は、Essay のタブを選んで、フィールドに入力すると文法リスト、CEFR レベルの表示と一緒に、脱落・誤形成・余剰の3つのタイプに色分けされたエラーチェックの結果が出力される（投野, 2023）。

ここで、先行研究で用いられていた DeepL と ProWritingAid を用いずに、DeepL Write と English Level Checker を分析用ツールとして利用したことについて補足しておきたい。DeepL は翻訳ツールで、Gold (2023) では調査対象者が未知語を調べたり、翻訳を行うために用いられていた。本研究ではこれらの機能を使用する必要がなく、文章を即座に推敲する機能を備えている DeepL Write の方が、AI ツールに英文の添削をさせるという観点からふさわしいと考え分析ツールに採用した。ProWritingAid は調査開始時に入手困難だったため、採用を見送った。その代わりに English Level Checker を採用した。English Level Checker は日本人研究者が開発しており、日本人英語学習者にとっての利便性があり、容易にアクセスが可能である点を考慮して分析ツールに選んだ。

なお、ChatGPT はプロンプト（指令文）を用いるという点で他の3つのツールと異なっており、添削結果に違いが見られる可能性があるが、今日急速に普及し、

幅広い分野で関心が高く、利用されている点、学習者が今後利用する機会が大いに高まる可能性を考慮して、分析ツールに選んだ。

3.3 データ

データは神奈川県内の公立高校で、筆者が指導するコミュニケーション英語Ⅱを受講する高校2年生が書いた学年末試験で書いた英作文とした。英作文は、海外からの留学生を短期間ホストファミリーとして受け入れる場合、神奈川県内で案内したい地名（場所）を1つ考え、案内する理由及び具体例（できることなど）を示した上で、25語以上40語以内の英語で書くものであった。試験は3クラスの118人が受験し、未解答及び語数不足のデータを除くと57人のデータを収集することができた。57人が書いた総語数は1,780語であった。入力語数が多すぎると分析できないツールがあるため、57人が書いた英文のうち、10人前後をひとまとめとして各ツールに分析させた。

当該の生徒には、調査の目的及びプライバシーに配慮することを伝え、データ利用の許可を得た。英検の取得状況は、2級が18人、準2級が8人、3級が19人、未取得が12人であった。作文には個人情報やプライバシーに関する情報が含まれていないことを確認し、筆者がワード文書に英文を入力し、個人が特定されないように、アルファベットと数字に置き換えて保存した。仮に1組1番の生徒はA-1、2組2番の生徒はB-2、3組3番の生徒はC-3とし、当該の生徒が書いた英文の直後に(A-1)や(B-2)などのようにそれぞれ表記して保存した。

3.4 分析手順

ChatGPT、DeepL Write、English Level Checker、Grammarlyの4つのライティングツールを用いてデータの分析を行った。DeepL Write及びEnglish Level Checkerにはそれぞれ入力欄にデータを入力し、出力結果をプリントスクリーン機能で複写し、パワーポイントのスライドに貼り付けてそれぞれ保存した。Grammarlyではデータをアップロードし、修正案が提示された画面をプリントスクリーン機能で複写し、パワーポイントに貼り付けて保存した。ChatGPTのバージョンはGPT 3.5を用い、Schmidt-Fajlik (2023)のプロンプト(What grammatical, spelling, or vocabulary mistakes are in the following text?)を参考に、次のようなプロンプトを作成した。

What grammatical, spelling, or vocabulary mistakes are in the following texts? Could you show me the original and corrected versions? Could you give me the reason why

you have corrected each error?

このプロンプトとデータを入力し、生成された回答を複写し、ワード文書に貼り付けて保存した。

各データを保存後、エクセルシート、元の英文と各ツールが示した修正箇所、変更箇所を見て、「修正案有」と「修正案無」の2つに区分し、修正内容を入力した。

「修正案有」は修正箇所について修正案が示されている場合を指す。例えば、元の英文がAlsoだけであったのに、Alsoの後にカンマが追加されているような場合は「カンマの追加」、itemがitemsのように複数形に修正されている場合は「複数形」、Yokohama stationがYokohama Stationのように修正されている場合は「大文字」のように入力した。“It is scared and funny.”(A-5)が“It is scary and fun.”のようにscaryとfunになっている場合は1箇所ずつ「語の修正」とし、エクセルシートに入力した。

「修正案無」は、修正案が示されず、語句等が修正されている場合を指す。例えば、“I want to go to My grandparents house with study abroad student. Because, my grandmother is good at make bread.”(A-15)をChatGPTは“I want to go to my grandparents’ house with a study abroad student because my grandmother is good at making bread.”のように修正している。元の英文の2文目はBecause、で始まっているが、修正された英文ではbecauseを従属接続詞と解釈し、前文と結びつけて1文で書き換えている。この場合何も言及されていないので、「修正案無」に分類した。

入力終了段階で、1つの英文に対して各ツールが示している修正案が一致している場合の一致数を調べた。1つの箇所に4つのツールすべてで一致している場合は一致数を4、3つのツールが一致している場合は3、2つの場合は2、1つの場合は1とした。例えば、“I want to tell you Akarenga.”(A-4)に対して、“I want to tell you about Akarenga.”とDeepL Write、English Level Checker、ChatGPTはaboutを修正案として示しており、一致数を3とした。

ChatGPTは“‘introduce Yokohama’ was retained for clarity.”のように提案はしたものの修正していない場合が見られ、この場合はデータに含めなかった。

4. 結果

57人が書いた英文に対して、4つのツールで何らかの指摘があった箇所は分析ツール間の重複を含めて合計で623箇所であった。表1は、ツール毎に修正案が示された場合と修正案が示されなかった場合について表したものである。

表 1. ツール毎の修正案の有無

ツール名	DeepL Write	English Level Checker	Grammarly	ChatGPT
修正案有	221	211	185	273
修正案無	10	0	0	81

最も修正箇所を示したのは ChatGPT の 273 箇所、次に DeepL Write の 221 箇所、English Level Checker の 211 箇所、最も少なかったのは Grammarly の 185 箇所であった。ChatGPT は 81 箇所、DeepL Write は 10 箇所、English Level Checker 及び Grammarly は修正案を示さずに修正することはなかった。

次に 4 つのツールで示された修正案有の修正箇所の一致数と箇所の数についてみていく。表 2 はこれをまとめたものである。

表 2. 修正案有の一致数と箇所数

一致数	箇所数
4	40
3	61
2	90
1	369
合計	560

4 つのツールで修正案が一致した箇所は 40 箇所、3 つのツールで修正案が一致したのは 61 箇所、2 つのツールは 90 箇所、1 つのツールが示した修正案は 369 箇所であった。

次からは、ツールの修正案有の一致数毎にみていく。表 3 は 4 つのツールで修正案が一致したときの内訳を表したものである。

表 3. 一致数 4 のときの内訳

内容	一致数	比率(%)
大文字	1	2.5
語の修正	8	20.0
語の追加	7	17.5
単数形	1	2.5
綴りの修正	14	35.0
複数形	9	22.5
合計	40	100.0

4 つのツールで最も一致したのは「綴りの修正」で 14 箇所 (35.0%)、次に「複数形」が 9 箇所 (22.5%)、

「語の修正」が 8 箇所 (20.0%)、「語の追加」が 7 箇所 (17.5%)、「大文字」と「単数形」が 1 箇所 (2.5%) ずつとなっていた。

「綴りの修正」では *recomend* や *beautiful, delisious, introdce, paticular* などの語が正しい形でそれぞれ修正案として示されていた。単純な綴りの誤りは 4 つのツールとも修正案を示せることがわかる。「複数形」では、例えば “If we go to Kamakura, we can see a lot of old building.” (B-18) の *building* を *a lot of* の修飾語句から *buildings* のように複数形に直す必要があるものから、“I would like to introduce Yokohama to overseas student.” (A-1) のように文脈から *student* を *students* に修正するものまで見られた。「語の修正」では、“There are many interesting shops, so I would like to buy something to the international student.” (A-8) の *buy* に対応する *for* を 4 つのツールで “I recomend their to go to Enoshima beause there are many foods. Also, there is a suizokukan. It has a lot of fish to watch. So I recomend their.” (B-10) では、1 文目と 3 文目の *their* を *them* に 4 つのツールすべてが修正していた。なお、*their* から *them* への「語の修正」はそれぞれ 1 箇所としてカウントした。「語の追加」では、例えば “I want to guide them to Yokohama. because Yokohama has lot of place to play.” (A-9) の *lot* の前の冠詞の *a* や “I want go Enoshima with foreign student.” (C-25) では *go* の前後にそれぞれ *to* を追加するといったようなものがあった。「大文字」については、“My favorite place is Landmark tower.” (C-2) を *Tower* への修正案、「単数形」については “Because They can learn Japanese cultures there.” (A-12) の *cultures* を *culture* に修正する案だった。

次に 3 つのツールで一致していた内容についてみていく。表 4 は 3 つのツールで修正案が一致したときの内訳を表したものである。

表 4. 一致数 3 のときの内訳

内容	一致数	比率(%)
大文字	5	8.2
カンマの追加	2	3.3
語句の修正	1	1.6
語の削除	4	6.6
語の修正	17	27.9
語の追加	17	27.9
小文字	3	4.9
単数形	1	1.6
綴りの修正	7	11.5
綴りの修正+	1	1.6
複数形		
複数形	3	4.9

合計	61	100.0
----	----	-------

3つのツールで最も一致したのは「語の修正」と「語の追加」が17箇所(27.9%)ずつで、次いで「綴りの修正」が7箇所(11.5%)、「大文字」が5箇所(8.2%)、「小文字」、「複数形」が3箇所(4.9%)ずつ、「カンマの追加」が2箇所(3.3%)、「語句の修正」、「単数形」、「綴りの修正+複数形」が1箇所(1.6%)ずつであった。4つのツールで最も一致していた「綴りの修正」は7箇所となったが、上位から3番目の値を示している。これまで触れなかった「語の削除」、「語句の修正」、「綴りの修正+複数形」の具体例を示しておきたい。「語の削除」については、例えば“Because The Kamakura is a lot of place where feel Japanese culture.”(B-20)では、DeepL Writeを除く3つのツールでTheを削除し、Because Kamakuraと修正案が示されていた。「語句の修正」は、“May be it makes you happy.”(B-2)のMay beをDeepL Writeを除く3つのツールでMaybeの1語に修正していた。「綴りの修正+複数形」は“and we can ride many attraction in kosumoworld.”(A-5)のattractionについてDeepL Writeを除く3つのツールでattractionsと修正されていた。3つのツールが修正したのにDeepL Writeだけ修正できなかったことについては考察の中で触れたい。

ここで3つのツールの組み合わせによる修正案有で修正箇所の一致数を表5にまとめてみた。

表5. 3つのツールの組み合わせによる一致数

ツール名	一致数	比率(%)
D, E, G	14	23.0
D, E, C	14	23.0
D, G, C	9	14.7
E, G, C	24	39.3
合計	61	100.0

注. DはDeepL Write, EはEnglish Level Checker, GはGrammarly, CはChatGPTを表す。

English Level CheckerとGrammarlyとChatGPTによる一致数が24箇所(39.3%)で最も多かった。次に一致数が多かったのは、DeepL WriteとEnglish Level CheckerとGrammarlyの組み合わせとDeepL WriteとEnglish Level CheckerとChatGPTの組み合わせの14箇所(23.0%)ずつ、DeepL WriteとGrammarlyとChatGPTの9箇所(14.7%)であった。

表6に3つのツールの組み合わせ毎の内訳をまとめた。

表6. 3つのツールの組み合わせ毎の内訳

3つのツール名	D, E, G	D, E, C	D, G, C	E, G, C	計
大文字	2	0	3	0	5
カンマの追加	1	0	1	0	2
語句の修正	0	0	0	1	1
語の削除	0	1	0	3	4
語の修正	5	4	1	7	17
語の追加	4	6	0	7	17
小文字	0	3	0	0	3
単数形	0	0	1	0	1
綴りの修正+複数形	0	0	0	1	1
綴りの修正	1	0	2	4	7
複数形	1	0	1	1	3
合計	14	14	9	24	61

注. DはDeepL Write, EはEnglish Level Checker, GはGrammarly, CはChatGPTを表す。

English Level CheckerとGrammarlyとChatGPTの組み合わせが24箇所でも最も一致数が多い。これは「0」の数値が4項目だけで最も少なく、「語の修正」と「語の追加」が7箇所ずつで他の組み合わせより多いことが理由として考えられる。DeepL WriteとGrammarlyとChatGPTの組み合わせでの一致数が9箇所と最も少なくなっているが、「語の修正」、「語の追加」の項目が他の組み合わせと比較すると少ないことが原因とみられる。

次に2つのツールで一致していた内容についてみていく。表7は2つのツールで修正案が一致したときの内訳を表したものである。

表7. 一致数2のときの内訳

内容	一致数	比率(%)
大文字	4	4.4
カンマの追加	3	3.3
語句の修正	5	5.6
語句の追加	3	3.3
語の位置の修正	1	1.1
語の削除	10	11.1
語の修正	18	20.0

語の追加	27	30.0
語の追加+複数形	1	1.1
小文字	3	3.3
単数形	1	1.1
綴りの修正	6	6.7
綴りの修正+大文字	1	1.1
複数形	7	7.8
合計	90	100.0

「語の追加」が 27 箇所 (30.0%) で最も多い一致数であった。以下、「語の修正」が 18 か所 (20.0%)、「語の削除」が 10 か所 (11.1%)、「複数形」が 7 箇所 (7.8%) などと続いている。これまでには見られなかった「語句の追加」、「語の位置の修正」、「語の追加+複数形」、「綴りの修正+大文字」といった項目が追加されている。「語句の追加」は、例えば DeepL Write と ChatGPT が “I have two reasons. First, many temples in Kamakura.” (C-22) に対して “First, there are many temples in Kamakura.” と many temples の前に there are を追加する案を示している。「語の位置の修正」は “And I want to enjoy air to clean in Hakone!” (C-20) で、DeepL Write と ChatGPT が “enjoy the clean air” のように修正案を示していた。「語の追加+複数形」は “In fall, the color of tree change for red, orange or yellow.” (A-23) で、DeepL Write と ChatGPT が tree を the trees と修正案を示していた。「綴りの修正+大文字」は、 “For example, Yokohama chukagai, Raund one and Cosumo wald.” (A-37) に対して、DeepL Write と ChatGPT は Raund one を Round One のように修正案を示していた。

ここで 2 つのツールの組み合わせによる修正案有で修正箇所が一致数を表 8 にまとめてみた。

表 8. 一致数 2 のときの内訳

ツール名	一致数	比率(%)
D, E	14	15.5
D, G	8	8.9
D, C	18	20.0
E, G	24	26.7
E, C	16	17.8
G, C	10	11.1
合計	90	100.0

注. D は DeepL Write, E は English Level Checker, G は Grammarly, C は ChatGPT を表す。

2 つのツールの組み合わせでは English Level Checker と Grammarly の組み合わせが、24 箇所 (26.7%) で最も一致数が多く、次は DeepL Write と ChatGPT の 18

箇所 (20.0%)、English Level Checker と ChatGPT の 16 箇所 (17.8%)、DeepL Write と English Level Checker の 14 箇所 (15.5%)、Grammarly と ChatGPT の 10 箇所 (11.1%)、DeepL Write と Grammarly の 8 箇所 (8.9%) の順であった。

表 9 に 2 つのツールの組み合わせ毎の内訳をまとめた。

表 9. 2 つのツールの組み合わせ毎の内訳

2 つのツール名	D, E	D, G	D, C	E, G	E, C	G, C	計
大文字	1	1	1	1	0	0	4
カンマの追加	0	3	0	0	0	0	3
語句の修正	0	0	2	1	1	1	5
語句の追加	0	0	1	0	2	0	3
語の位置の修正	0	0	1	0	0	0	1
語の削除	0	0	2	4	4	0	10
語の修正	4	0	3	6	4	1	18
語の追加	6	2	4	8	5	2	27
語の追加+複数形	0	0	1	0	0	0	1
小文字	1	0	2	0	0	0	3
単数形	0	1	0	0	0	0	1
綴りの修正	1	1	1	0	0	3	6
綴りの修正+大文字	1	0	0	0	0	0	1
複数形	0	0	0	4	0	3	7
合計	14	8	18	24	16	10	90

注. D は DeepL Write, E は English Level Checker, G は Grammarly, C は ChatGPT を表す。

DeepL Write と Grammarly の組み合わせの一致箇所は 8 か所で最も少ないが、他の組み合わせで見られない「カンマの追加」が 3 箇所一致している。DeepL Write と ChatGPT の組み合わせは 18 箇所 2 番目に多

いが、English Level Checker と Grammarly の 24 箇所よりも「0」の項目が少なく、幅広い項目で一致数が多く見られる。

「語の修正」は 18 箇所、2 番目に一致数が多いが、DeepL Write と Grammarly の組み合わせのみ「0」となっている。「語の修正」で最も多いのは English Level Checker と Grammarly の組み合わせの「6」となっているが、例えば“*And Komatidori is a lot of shop.*” (B-20) で English Level Checker と Grammarly が *is* を *has* に修正するように、また ChatGPT は修正案を示さずに *has* に修正している。しかし DeepL Write のみ、この英文に何も修正を入れていない。Komatidori が小町通であることを DeepL Write は判断できなかっただけでなく、*a lot of* があるにもかかわらず *shops* を複数形にも修正していない。ただし、*a lot of shop* に関しては Grammarly が *shops* に、ChatGPT は *many shops* を修正案として提示しているが、English Level Checker は *shop* のままであった。

「語の修正」の前後にさらに「語の削除」などがある場合に「語の修正」が行われていないこともあった。

一致数が 2 でも引き続き「綴りの修正」も見られるが、Grammarly と ChatGPT で固有名詞 (Yamasita を Yamashita に、CosmoWorld を Cosmo World) に修正、*dericious* を *delicious* に修正するものが見られた。*dericious* については DeepL Write が *delicacies* に English Level Checker が「語の削除」を提案しているところが興味深い。

1 つのツールだけで修正案を示したのは、369 箇所であった。「語の修正」や「語句の修正」が多く見られる中で、「綴りの修正」も見られた。ChatGPT は固有名詞や特定の産物に威力を発揮するようで、Komatidori を Komachi-dori、Kamakuradaibutu を Kamakura Daibutsu、shouronpo を xiaolongbao に直すなど、他のツールとは異なる側面を見せた。また “*I intrudece Enoshima.*”

(B-35) については ChatGPT のみが修正案を示し、他のツールは修正していなかった。

「修正案無」は ChatGPT が 81 箇所、で、DeepL Write は 10 箇所、で修正案を示さずに修正していた。ChatGPT で特に目立つのは *because* の扱いである。例えば “*I want to introduce to Yokohama. Because Yokohama has very beautiful night view.*” (B-39) を他の 3 つのツールでは *Because* で始まる第 2 文を修正せずにこのままにし、その他の箇所について修正案を示しているが、ChatGPT は “*I want to introduce you to Yokohama because Yokohama has a very beautiful night view.*” のように修正していた。ただし、前文に続けて *because* を接続詞として修正したことについての修正案は示されなかった。81 箇所中 17 箇所も *because* に修正する実態が見られた。

このことについては考察の中で触れる。

5. 考察

57 人の高校生が書いた英作文データを ChatGPT、DeepL Write、English Level Checker、Grammarly の 4 つのツールで分析を行った結果、何らかの指摘があった箇所は分析ツール間の重複を含めて合計で 623 箇所であった。「修正案有」に関して最も修正箇所を示したのは ChatGPT の 273 箇所、次に DeepL Write の 221 箇所、English Level Checker の 211 箇所、Grammarly の 185 箇所であった。「修正案無」に関しては ChatGPT が 81 箇所、DeepL Write は 10 箇所、English Level Checker 及び Grammarly は 0 箇所であった。ChatGPT が修正案の有無にかかわらず最も修正し、DeepL Write も ChatGPT には及ばないものの修正が多いことがわかる。

4 つのツールで修正案が一致した箇所は 40 箇所、3 つのツールでは 61 箇所、2 つのツールは 90 箇所、1 つのツールが示した修正案は 369 箇所であった。ツール数が増えるにつれて、修正案の一致数は減っている。

4 つのツールで最も一致したのは「綴りの修正」が 14 箇所 (35.0%)、次に「複数形」が 9 箇所 (22.5%)、「語の修正」が 8 箇所 (20.0%)、「語の追加」が 7 箇所 (17.5%)、「大文字」と「単数形」が 1 箇所 (2.5%) ずつとなっていた。「綴りの修正」が 4 つのツールで最も一致していたが、*recomend* や *beautifull* のような単純な綴りの誤りは 4 つのツールとも修正案を提示できるということが判明した。

3 つのツールの修正案は 61 箇所、最も一致したのは「語の修正」と「語の追加」が 17 箇所 (27.9%) ずつ、次いで「綴りの修正」が 7 箇所 (11.5%)、「大文字」が 5 箇所 (8.2%)、「小文字」が 3 箇所 (4.9%)、「カンマの追加」が 2 箇所 (3.3%)、「語句の修正」、「単数形」、「綴りの修正+複数形」が 1 箇所 (1.6%) ずつであった。ここで、なぜ 4 つのツールでの一致が見られなかったのか検討してみたい。「綴りの修正」は 4 つのツールで最も一致数が多かったが、3 つのツールでは 7 箇所であった。2 つの例を取り上げたい。まず 1 つ目として、“*For example, Ebina are good food and big shooping store.*” (B-14) を取り上げる。この英文では、DeepL Write のみ *shooping* を容認し、他の 3 つのツールは *shopping* の修正案を提示している。この生徒は次の英文で “*They always shopping.*” と書いており、文法的に誤った英文を産出しているが *shopping* を用いているので、*shooping* を *shopping* と修正しても良かったはずだが何も修正されていなかった。2 つ目として “*For example, you can eat ra-men there.*” (A-24) を取り上げる。この英文の *ra-men* を DeepL Write、Grammarly、

ChatGPT は ramen と修正案を示している。ChatGPT に至っては、“‘ra-men’ was corrected to ‘ramen’ for proper spelling and clarity.”と明確に示している。英英辞典でも ramen は採用されて、英語として認知されているにもかかわらず、1つのツールで認識されなかったため、3つの一致にとどまったものと考えられる。Deep L Write と English Level Checker と ChatGPT の組み合わせの時には「綴りの修正」が見られなかったことから Grammarly は「綴りの修正」に優れている可能性がある。

ここからは前節で触れた「語の削除」、「語句の修正」、「綴りの修正+複数形」からみていく。「語の削除」は、“Because The Kamakura is a lot of place where feel Japanese culture.” (B-20) で見られたが、DeepL Write を除く3つのツールで The を削除し、Because Kamakura と修正案が示されていた。この生徒が書いた英文はその直前の英文も示すと“I introduce the Kamakura to foreign students. Because The Kamakura is a lot of place where feel Japanese culture.”であった。DeepL Write はここで I introduce Kamakura と修正案を提示せずに修正していた。次の英文では Kamakura の直前にあった The を the に修正するように案を提示していた。文の途中であると判断して、小文字で始まる the という修正案を出したと考えられるが、その前文では the を削除しているのが不可解である。

「語句の修正」は、“May be it makes you happy.” (B-2) の May be を DeepL Write を除く3つのツールで Maybe の1語に修正していた。DeepL Write はこの英文を “It might make you happy.”とし、It might make を修正案として提示している。このため4つのツールすべてで一致が見られなかったと考えられる。

「綴りの修正+複数形」は “And we can ride many attraction in kosumoworld.” (A-5) の attraction について DeepL Write を除く3つのツールで attractions と修正されていた。ChatGPT はこの部分を “...views, and there are many attractions in Cosmo World” と修正し、“many attractions in Cosmo World” を修正案として提示している。ただし、その直前の “, and there are” について修正案は示していない。日本語で「アトラクションに乗る」をそのまま英語にしても表現できないことから、ChatGPT は語句を修正し、DeepL Write は attraction を何か特別な乗り物と解釈し、attraction とは解釈しなかった可能性がある。

この他に3つのツールで共通していても1つのツールが別の修正案を提示している場合が見られる。例えば、「語の修正」では “I want to suppose to Noge-yama zoo.” (A-27) では、DeepL Write、English Level Checker、Grammarly が suppose の代わりに go を修正案とし、

ChatGPT は suppose to の代わりに visit を修正案として提示している。単純な修正ならば go で十分かもしれないが、visit のように修正案を示すことで、学習者が批判的に熟慮し賢明に判断できる機会を与えることなる可能性がある。

「語の追加」でも同様な傾向が見られた。“I want to guide the family to Hakone. There are one of the most famous hot springs there. So I want to go to the hot spring and eat a black egg with host family.” (B-3) では、English Level Checker、Grammarly、ChatGPT が最終文の host family の直前に my を修正案として提示しているのに対して、DeepL Write のみが the を修正案として提示している。第1文に the family が見られるので、the host family と修正したのかもしれない。ここでは my の方が文脈からふさわしいように思われる。学習者にこれらのツールの修正案を提示し、どちらの修正がふさわしいか熟慮させることで賢明な判断を下せる機会を提供することになるだろう。

3つのツールの組み合わせでは、English Level Checker と Grammarly と ChatGPT の組み合わせが24箇所でも最も一致数が多かった一方で、DeepL Write と Grammarly と ChatGPT の組み合わせは9箇所と最も少なかった。少なかった要因として、「語の修正」、「語の追加」の項目が他の組み合わせと比較すると少なかったことが考えられる。

2つのツールでは「語の追加」が27箇所(30.0%)で最も多い一致数であった。以下、「語の修正」が18か所(20.0%)、「語の削除」が10か所(11.1%)、「複数形」が7箇所(7.8%)などと続いていた。2つのツールで一致する項目は3つのツールで一致する項目よりも増えていた。「語句の追加」、「語の位置の修正」、「語の追加+複数形」、「綴りの修正+大文字」などの修正案から、1箇所につき複数の修正案が提示されていることが見てくる。

ここからはいくつかの興味深い提案について触れてみたい。1つのツールのみが修正案を提案している中で興味深いものが見られた。“I want to go to Enoshima with study abroad students.” (A-34) に関して、DeepL Write のみが abroad を foreign と修正するように提案し、他の3つのツールは何ら修正を施していなかった。

複数の英文でも興味深い現象が見られた。例えば、“I would like to go to Kamakura with foreign student. I have two reasons. First, many temples in Kamakura. We feel Japanese culture. Second, we can eat many delicious food. So, I want to go to Kamakura with foreign student.” (C-22) をとりあげる。1文目の foreign student に関して、DeepL Write、English Level Checker、

ChatGPT は冠詞の a を追加するように提案しているのに対して、Grammarly は students と複数形にするように提案している。第 3 文では many の直前に there are を追加するように DeepL Write と ChatGPT は提案しているが他の 2 つのツールは何ら言及がない。第 4 文については、DeepL Write が We can feel と can を追加しているが、ChatGPT は第 4 文を第 3 文に続けて、“... Kamakura, and we can experience Japanese culture.”と語句の修正を提案している。第 5 文については ChatGPT が eat を enjoy に「語の修正」をするように、food を foods の複数形にするよう提案している。DeepL Write は many に代わって a lot of を修正案として示しているのに対して、English Level Checker は many を削除するように提案している。最終文については、English Level Checker と Grammarly が student を複数形にするように提案し、ChatGPT は student の前に冠詞の a を追加するように提案している。学習者はここでどのツールの修正案が最もふさわしいか熟慮し、修正していくことになるが、習熟度の低い学習者には的確に判断するのが難しい可能性がある。

1 人が書いた英作文の中で 4 つのツールで異なる修正案を示しているものがあった。例えば、“I want to suppose to NogeYama zoo. I have two reasons. First, NogeYama zoo is free. It is convenient to see animals. Second, we can see many animals. It may enjoy there.”

(A-27) の最終文では、DeepL Write が may を can に修正するように、English Level Checker が enjoy の後に it を追加するように、Grammarly が may と enjoy の間に be を追加するように、ChatGPT は “and it can be an enjoyable experience” と語句を修正するように提案し、解釈が異なっていた。ここでは ChatGPT の修正案が文脈に最も適切な修正案を提示していると考えられるが、学習者が分析力を培うにはふさわしい例かもしれない。

「修正案無」で ChatGPT が 81 箇所中 17 箇所について Because で始まる主節を欠いた英文を、because として前文に続けて 1 文に修正する英文を提示した。この場合、本来ならば修正案を提示するところだが、ChatGPT は修正案を提示することはなかった。主節を欠いた「Because+主語+動詞」について、英和辞典に次のように記載されている。

[Because で文を始めない]よく日本人は I can't go to see the baseball game on Saturday. Because I have to study. のように主節と because 節を別々に書くが、これは誤り。I can't go to see the baseball game on Saturday because I have to study. (土曜日に野球の試合を見に行くことはできません。勉強しなければならぬからです。) のようにするのが正しい。た

だし、Why 疑問文に答える場合は別。(南出・中邑, 2023, p. 182) (強調は原文のまま)

主節のない「Because+主語+動詞」の英文は、Why で始まる質問文の応答及び躊躇して言葉をつなぐ場合に見られる (Swan, 2016) が、本来 because は接続詞なので、南出・中邑 (2023) が述べるように記す必要がある。今回のデータの中で、ChatGPT が because に書き換えなかったのは 1 例だけ見られたが、修正案を示さずに 17 例は確実に修正されていた。他の 3 つのライティングツールでは主節を欠く Because で始まる英文を容認していた。この点に関しては、現段階の AI を活用したライティングツールに依存すると学習者は正用法を身に付けることができない可能性があり、指導者が授業や添削の中で丁寧に指導していく必要があるだろう。

6. おわりに

本研究では、高校生が書いた英作文を複数の AI を活用したライティングツールに添削させ、ライティングツールが示す修正案を含む修正結果の一致点、一致しない点を比較、分析、考察し、今後の英作文指導に資することを目的として 3 つの研究課題を設定した。研究課題 1 は「複数のライティングツールが示す修正案はどの程度一致が見られるか。」であった。4 つのツールで修正案が一致した箇所は 40 箇所、3 つのツールでは 61 箇所、2 つのツールは 90 箇所、1 つのツールが示した修正案は 369 箇所であった。ツール数が増えるにつれて、修正案の一致数は減っていることが判明した。この結果から、習熟度の高い学習者には、1 つのツールよりも複数のツールの利用を勧めるのも一考に値すると思われる。4 つのツールがすべて一致して指摘する修正案は、学習者が自信を持って修正しやすい可能性がある。また複数のツールで一致数が見られない箇所については学習者が批判的に熟慮し、賢明に判断する必要が生じるだろう。どのツールが適切な修正案を提示しているのか、あるいはいないのかを吟味することで、自己編集力や正確性が向上する可能性がある。この場合、必要に応じて指導者が関与していくことも考えられる。

研究課題 2 は「複数のライティングツールが示す修正案にはどのような内容が見られるか。」であった。4 つのツールでは「綴りの修正」が最も一致し、「語の修正」、「複数形」、「語の追加」、「大文字」と「単数形」の順で一致数が減っていった。3 つのツールではこれらに加えて、「小文字」、「カンマの追加」、「語句の修正」、「単数形」、「綴りの修正+複数形」といった修正案が見られた。2 つのツールではさらに「語句の追加」、「語

の位置の修正」、「語の追加+複数形」、「綴りの修正+大文字」といった項目が追加されており、ツールの一致数が少なくなるにつれて修正案の内容が増加する傾向にあることがわかる。ツール数が増えるにつれて、修正案の内容が増加することで学習者は修正内容を一つひとつ慎重に見極める必要が生じるので、習熟度の低い学習者には指導者の支援が必要となるだろう。ただし、「綴りの修正」は4つのツールとも修正案を提示する傾向にあることから、英作文の指導ではどのツールを勧めても構わないように思われる。特に、「綴りの修正」に関しては、今回調査した4つのツールの中で、Grammarlyが優れている可能性が高く、学習者に使用させ実感させてみると良いかもしれない。

研究課題3は「複数のライティングツールの中には、修正案を示さずに修正している部分は見られるか。」であった。「修正案無」でChatGPTが81箇所、DeepL Writeが10箇所で修正案を示さずに修正していたが、English Level CheckerやGrammarlyではこのようなことは見られなかった。修正案を示さずに英文が修正されている場合、習熟度の低い学習者はそのことに気づかずに、修正案が示された箇所だけを検討し修正する可能性が高くなる。教員が英作文でライティングツールを用いて添削指導を行う場合、修正案が提示されていないのに、修正されている部分がないかどうか、学習者への気づきを促しながら、指導していく必要があるだろう。

特に文頭のBecauseを大文字で始めて主節を持たない英文に関しては、ChatGPT以外のツールは修正を施すことはなかった。ChatGPTは他のツールよりも修正箇所が多いが、修正した箇所についての「修正案」を提示されないことが多々見られたので、習熟度の低い学習者にとっては丁寧に指導する必要があるだろう。これはChatGPTのプロンプトを工夫することで解決できる可能性がある。他の3つのライティングツールはプロンプトがなく、かつ容易に利用できてしまうので、「修正案」の有無に関わらず、学習者自身が書いた元の英文とライティングツールが提示した英文を比較・検討させる必要がある。

本研究では学習者に4つのツールの結果をフィードバックし、実際に修正を考えさせるところまでは至らなかった。また、今回訂正フィードバックの一致率について調査したが、提示された修正案のうち、正しい誤りの訂正がどの程度の割合を占めていたのかまでは示すことができなかった。さらに、ライティングツールで修正案が提示された誤りとして、全体的な誤り(global error)と局所的な誤り(local error)に当たるものの程度についても調査できなかった。こういった点を追求することで、ライティングツールの選定や使用方

法、そしてツールの抱える課題がより鮮明になる可能性がある。今後の研究課題として取り組んでいきたい。

生成AIを活用したライティングツールは進化の途上にあり、今後さらに精度が高まることが予想される。学習者自身が有する知識や経験に照らし合わせて、ライティングツールの提案を受け入れるかどうかを判断し、決断するのは学習者である人間であることを学習者に伝えたい。

謝辞

本稿の完成にあたり、貴重なご意見をくださった3名の匿名査読者及び日本英語教育学会・日本教育言語学会合同編集委員会に、この場をお借りして厚く御礼申し上げます。

文献

- Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4), 584-607.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102-118.
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14(3), 191-205.
- DeepL. (2022). DeepL Write. <https://support.deepl.com/hc/ja/articles/6318834492700-DeepL-Write> (2023年10月20日閲覧)
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353-371.
- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8(1), 1-11.
- Ferris, D. (2002). *Treatment of Error in Second Language Student Writing*. University of Michigan Press.
- Gold, J. (2023). Fostering self-directed writers using AWE and WBMT: A pilot case study. *Journal of the School of Languages and Communication, Kobe University*, 19, 1-27.
- Hyland, F., Nicolás-Conesa, F., & Cerezo, L. (2018). Key issues of debate about feedback on writing. In R. M. Manchón, & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 433-452),

Mouton De Gruyter.

- 南出康世・中邑光男（編集主幹）（2023）『ジーニアス英和辞典 第6版』大修館書店.
- 水本篤（2023）「生成系 AI は英語教育研究をどう変えるのか？」『英語教育』72(6), 61-63.
- 文部科学省（2023）「初等中等教育段階における生成 AI の利用に関する暫定的なガイドライン」
https://www.mext.go.jp/a_menu/other/mext_02412.html（2023年10月20日閲覧）
- 新美徳康・梅木璃子（2023）「自動ライティング評価ツール Grammarly を援用した英語ライティング指導の探究的実践」『リメディアル教育研究』.
- Schmidt-Fajlik, R. (2023). ChatGPT as a grammar checker for Japanese English language learners: A comparison with Grammarly and ProWritingAid. *AsiaCALL Online Journal*, 14(1), 105-119.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255-283.
- Swan, M. (2016). *Practical English Usage* (4th ed.). Oxford University Press.
- 投野由紀夫（2023）「教材研究を深めるコーパス活用のツボ（5）：教科書・英作文の CEFR レベルをチェックする：English Level Checker」『英語教育』71(12), 60-61.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327-369.
- Truscott, J. (1999). The case for “The case against grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122.
- 柳瀬陽介（2023a）「AI を活用して英語論文を作成する日本語話者にとっての課題とその対策」『情報の科学と技術』73(6), 219-224.
- 柳瀬陽介（2023b）「ChatGPT は孫悟空」『英語教育』72(6), 58-60.