

# 自動添削および自動採点と機械翻訳を用いたライティング指導の可能性： 英文作成支援ツール Transable を使用して\*

築地原 尚 美  
山下 美 朋

## 1. はじめに

英語教育，特にライティング教育は大きな変化を迎えている。2016年にニューラル翻訳が開発され，TOEIC900点以上の英文が機械翻訳で簡単に書けるようになり（小田，2019；山田，2021），機械翻訳を使用してエッセイ課題を仕上げる生徒や学生が出てきた。また2022年11月にはChatGPTが開発され，生徒や学生がChatGPTに英文を書かせる可能性があるとの懸念から，高等教育の現場では，学生の生成AI使用についての指針を示している。更には，ライティング指導におけるフィードバックも，自動添削および自動採点ツール（Automated Writing Evaluation, AWE）の開発により，教員フィードバックと併用する指導が行われつつある（Mizumoto & Eguchi, 2023）。このような背景のもとで，筆者らは，高等学校または大学の英語教育の現場で，機械翻訳や自動採点ツールを用いたライティング指導を行なっ

---

\* 本稿は，2023年8月8日に早稲田大学にて開催された外国語教育メディア学会において研究発表した内容に，少なからず加筆・修正を施したものである。発表の内容に有益な質問，コメントをいただいた皆様に記して感謝申し上げます。また，本稿の執筆に際して大変貴重なご指摘・ご助言をいただいた編集委員長ならびに匿名の査読委員の先生方に感謝申し上げます。お忙しい中査読を引き受け，貴重な査読コメントをくださり，深く感謝する次第である。なお，本稿における不備や誤りは全て筆者の責任による。

てきた。そして、これらのツール使用には賛否両論あるが、利活用を促すことで生徒や学生が自律的に英文を書くことができるツールの開発を目指してきた。本稿では、自動添削および自動採点と機械翻訳の指導における有効性を検証するために、筆者らが開発に関わったライティング支援ツール Transable と、既存の自動採点ツール Write and Improve with Cambridge を比較する。そして Transable の有用性を検証することで、ライティング指導における自動添削および自動採点と機械翻訳の可能性を探る。

## 2. 先行研究

### 2.1. 機械翻訳の先行研究

機械翻訳を外国語学習に取り入れる試みは2016年以前からあったが (Niño, 2008), その機械翻訳利用は、機械翻訳が産出した文章を Bad model とする指導が中心であった。学習者が、機械翻訳を使って母語から外国語への翻訳を行い、機械翻訳の産出文の中に誤りを見つけ、修正することにより、学習を促すというものである (Niño, 2008)。しかし2016年以降ニューラル翻訳が開発されて以来、機械翻訳の質が向上し、機械翻訳の産出文から学びを得る Good model アプローチが可能になった。この方法では、英文を書く際に、まず原文を日本語で作り、一方で英文も作成する。そして、その英文を、日本語の文から機械翻訳に産出させた英文と比較する方法である。AI翻訳は文法的・語彙的に正確な文を産出するため、機械翻訳産出文を参照し、自分が書いた英文を修正することで学びを得るとされている (Lee, 2020; Tsai, 2019; 山田, 2021)。

また機械翻訳のその他の利点として、認知的、言語的、感情的観点からの議論がある (Lee, 2020)。機械翻訳を用いることで母語に頼ることができるため、第二外国語に変換する際に外国語から母語に翻訳する脳への認知負荷を軽減し、母語を使うことで自律的な学習を促進する。言語学的観点においては、機械翻訳は学習者の語彙・文法知識を補うため、読解とライティング活動を容易にし、最終的に言語学習を促進する (Niño, 2009)。また、情緒的な観点においては、言語不安を低下させ、モチベーションと自信を高め (Niño, 2008), 学習不安を軽減する (Niño, 2009) と報告されている。また、機械翻訳は、コンピュータ支援言語学習 (Computer-Assisted

Language Learning, CALL) として、学生が仲間の学生や遠隔の人たちとの協同学習や協同推敲が可能で、自律学習をも支援するとされている。このような利点により機械翻訳の需要は学習の場面で高まっている。しかし信頼性はまだ十分に確立されているとは言えない。様々にある批判の一つは、機械翻訳は誤訳が多く、誤った語彙、不正確な文法を産出することである (Niño, 2008, 2009)。そのため、教師は機械翻訳から学習成果を得るために、学生にどのように使わせるのが最適な方法であるのかを見極める必要がある。つまり機械翻訳がどのように学生の言語学習を促し、コンピュータ支援言語学習を促進できるのかを調査し、教育における機械翻訳使用の潜在的な利点とその危険性を認識することが重要である。

## 2.2. 自動添削および自動採点の先行研究

自動添削および自動採点はプロジェクト・エッセイ・グレード (Project Essay Grade, PEG) に始まる (Page, 1967)。このプロジェクトでは、重回帰分析を利用し、平均的な文の長さ、平均的なエッセイの長さ、前置詞の数、カンマの数など、測定可能なテキストの特性に基づいてエッセイのスコアを予測し、人間の評価者が与えたスコアとの相関を導き出した。しかし、PEGは表面的な構造にのみ焦点を当て、内容に関連する特徴に弱いという批判があった。その後、1990年代にはコンピュータの進歩により、より精密な自動採点システムが作られるようになり、人間の知的能力をコンピュータ・プログラミングでシミュレートするAIの一分野である自然言語処理技術 (Natural Linguistic Programing, NLP) が大きく進歩した。この発展により、表面的な言語的特徴だけでなく、より正確な採点が可能になった。1998年に Educational Testing Service (ETS) によって開発された e-rater は、そのような新しい自動採点システムの代表である。e-rater システムは、文の構造 (統語論)、単語の構造 (形態論)、意味 (意味論) の分析を可能にする新しい手法を採用しているのである (Burstein, Tetreault & Madnani, 2013)。そのため、e-rater の信頼性と妥当性が報告されており、代表的な大学入学テスト Graduate Record Examination (GRE) や Test of English as a Foreign Language (TOEFL) のような高難易度のテストでは、人間の評価者と組合せて使用されている (Mizumoto & Eguchi, 2023)。

また、ライティング指導においてフィードバックを行うことは必須とさ

れているが、主に文法・語彙・構成・内容において行われることが多い。ライティング指導の最終目的は、「自律的な書き手を育てること」(佐渡島・太田, 2013)であり、自分で書いた英文に対して自分で修正できる力を育てる必要がある。これは何度も指導を受けることにより習得できるものとされ、Ferris (2017) は、修正可能な間違い(文法・構文など)と修正不可能な間違い(表現・内容・構成など)のうち前者の間違いを自分で修正できるように方向づけることが肝要であると述べている。しかしながら、修正に必要なフィードバックを行うことは教師にとって負担が大きい。そのため研究者は、この作業をコンピュータと共有する技術を考案しようと試みてきた。そのなかで、自動添削機能を備えた自動ライティング評価(Automated Writing Evaluation, AWE)が、近年利用されつつある。自動添削機能は、形式的な特徴に関するフィードバックが自動化され、教師の負担を軽減し、教師は構成、論証、内容の展開などより高いレベルのフィードバックに時間を割くことができる。また、自動添削および自動採点システムの即時性や、個別にフィードバックを行い学生に対して何度も復習の機会を提供することは、学生の復習スキルを向上させる利点がある。一方で、このようなツールを使用することに懐疑的な研究者もいる。それらのツールが測定するライティング構成要素は、教室で評価されるものとは異なるからである(Condon, 2013)。しかし、Criterion (ETS, 2021)、ChatGPTのようなツールは、内容・構成といったライティング構成要素を添削するため、学習者が英文の質を向上させるのに役立つとの報告もある(Mizumoto & Eguchi, 2023)。

以上、ライティング指導における機械翻訳、自動添削・自動採点ツールの利用また問題点を述べてきた。しかしChatGPTなどの生成AI搭載型ツールを教員が授業に導入した例は限定的で、まだ教育効果も示されていない。このような状況下で筆者らは、自律的学習を可能とする生成AI機能の利点を活かした英文作成支援ツールTransableを開発した。このようなツールの開発は例を見ず、筆者らが期待するのは学生が教員の介入なしに英文を書き、修正し、英文の質が向上することである。

### 3. 研究の目的

本研究の目的は、今回新たに開発した英文作成支援ツール Transable (Tr) の有用性を探ることである。Trには (1) 生成AIによる添削、(2) 機械翻訳、(3) 自動採点システム、が搭載されているが、その有用性を探るために既存の Write and Improve with Cambridge (WIC) と比較する。WICには、(2) はないが (1) と (3) があり、大学教育の現場で広く使われていることから、学生が (1) と (3) の助けを借りて英文の質が向上することが認められている。本稿では、機械翻訳機能が既に搭載されている Tr 利用に焦点を置き、Tr を用いた英作文で 1. 英文の質は向上するか、2. 自動採点は学生の英文の質を適切に判断しているかの 2 点を、Tr と WIC を比較することで探る。また、3. 学生が WIC と Tr を使ってライティングをした際にそれらに対して感じた使用感を調査し、機械翻訳と自動添削および採点の可能性について議論する。

### 4. 方法

#### 4.1. 使用ツール

今回開発した Transable (<https://transable.net>) は、筆者らが考案のもと立命館大学の大学院生杉山滉平氏により具現化された英文作成支援ツールである。DeepL, Grammarly, ChatGPT を組み込むことで機械翻訳、自動添削および自動採点機能を搭載したツールとなっている。図 1 は Tr の画面の例である。学生が左の「エッセイを書く」というウィンドウ内に英文を書く

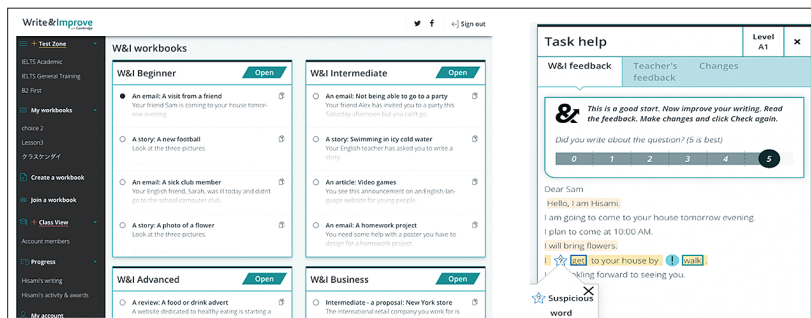
図 1 Transable (Tr) の画面例



と、修正すべき箇所はGrammarlyにより下線で示される。図1の「エッセイを書く」内の文を見ると、Grammarlyが指摘している文法・語法の誤りが赤色で示されている(itsやThereforeなど)。また、右の「表現を調べる」というウィンドウで知りたい単語や表現を調べたり、搭載されたDeepLで内容を日本語で確認することもできる。更には、英文を書き終えた時点で「評価する」を押すと、GTEC<sup>1</sup>とIELTS<sup>2</sup>、Common European Framework of Reference for Languages (CEFR)<sup>3</sup>のルーブリック評価を得ることができる。これはChatGPTによる評価である。筆者らが、学生のサンプルエッセイを何度も評価させることで、ChatGPTにより一定の評価が得られるように実験を重ねて作成した。

もうひとつのツールWIC (https://writeandimprove.com/) は、Cambridge大学、Cambridge大学検定機構、Cambridge大学出版が共同開発した自動採点ツールである。学生が書いた英文に対してAIが文法・語彙の誤りを指摘するライティング支援機能もついており(図2)、それらは主に文法・語彙・句読点の修正を指摘する機能である。例えば、黄色のハイライトは、その文中に文法の誤りがあることを示し、星印などの記号は、文中に何らかの語が入るという意味で、ともに修正を促している。黄色の斜線ハイライトで示された文は、もう少しふさわしい表現に書き換えることが可能なことを示している。学生は、これらのAIの指摘に基づき誤りを訂正したり、文章を書き直したりすることができる。そして書いた英文は、CEFRのルーブリックに基づいて自動採点してもらえ機能が付いている。学生が何度も書き直しをしながら英文を繰り返し自動添削することで、ライティング

図2 Write and Improve with Cambridge (WIC) の画面例



活動に臨むのを支援するツールである。

#### 4.2. 参加者

本研究の対象者は、滋賀県の公立大学非英語専攻理系2クラスの学生（計60名）であり、2023年4月から8月に行った授業の中でデータを収集した。母語は日本語であり、英語力はCEFRでlevel A2程度である。データの収集は倫理委員会の承認を得ており、同意を得た学生のもののみ収集した。

#### 4.3. 実践実験授業のタスクと手順

WICには機械翻訳機能はないため、学生には好みの機械翻訳や辞書を利用させた。グーグル翻訳やDeepLを使用する学生がいた一方で、文法書やオンライン辞書などを使う学生もいた。WICに対し、Trは先述の通り、DeepL, Grammarly, ChatGPTが搭載されている新規開発されたライティングツールである。本研究においては、既存の自動採点ツールで商用化されているWICを使わせる群と、Trを使わせる群の二つに分け調査を行った。二群に分けた理由は、この度新たに開発されたTrの性能とその有効性、また学生の英文の質の変化をWICと比較して見るためである。そして、学生には大学の英語Iの授業で用いている三修社の大学英語四技能錬成用テキスト「Develop your English with four skills」を用いて、下記のテーマのエッセイを三回書いてもらった。それぞれの英文のプロンプトは次の通りである。

1. Do you want to eat Japanese food while travelling abroad?
2. Will solar-powered flower beds installed in public have good effects on Japanese society?
3. Suppose you have a friend who wants to have plastic surgery. Do you support his/her decision?

各授業回での手順は次の通りであった。学生は上記のプロンプトを与えられた後、内容について数分間考え、英文を書いた。そして60名のうち30名はWIC、残りの30名はTrを使い、書いた英文を修正した。その後、図3が示すワークシートを作成した。ワークシートには自力で書いた英文と修正英文それぞれの総語数、異なり語数（土肥・須藤・中村, 2023）、語の訂正

数、表現の訂正数、文の訂正数及び修正箇所を色分けして記入させた。異なり語数とは、総語数から重なり語数を差し引いた語数のことであり、学生にカウントさせて記入させたものを、教員が再度点検した。更に自動採点機能を用いて修正前と修正後のCEFRのレベルを出し、記入させた。この手順は3つのプロンプトでそれぞれ行い、ワークシートは一人3枚提出した。また、TrとWICを使用した授業に関するアンケートを実施し回答させた。

#### 4.4. 分析方法

学生によって提出された3枚のワークシート（総数180枚うち提出数167枚、残り13枚は学生の欠席および課題未提出により欠損）に記載された自力で書いた英文と修正英文それぞれの「総語数」、「総語数の変化」、「語のバリエーション（異なり語数）」、「語の訂正数」、「表現の訂正数」、「文の訂正数」を分析対象とした。また、WICとTrの群それぞれで、各ツールの自動採点機能を利用して得られたCEFRのルーブリックに基づく採点結果も調査した。

図3 ライティング活動のためのワークシート

修正後の英文	語数	( 189 )	語
語のバリエーション数 (全体の語数から同じ語の繰り返しを除く語数)		( 103 )	語
(修正した箇所は色をつけてください 赤：単語の修正, 緑：表現, 青：センテンス)			
誤り率 (修正前の語数のうちの誤りの語数の割合)			%
<p>I agree with the ideas that solar-powered flower beds installed in public have good effects on Japanese society. I have three reasons. First, they <b>make</b> public facilities colorful. They are easy to manage and need less cost because they <b>water</b> flowers with solar energy. So, more public facilities may put flowers. If there are beautiful flowers, people relax and feel clean. That will make people easier to use <b>the facilities</b>. Second, <b>flower beds</b> show <b>the activity of the environment</b> to people. Using <b>solar power</b> and raising plants have <b>a good influence on the environment</b>. That activity will make thinking of some people changed. <b>More people assist action for nature, too</b>. A big organization's influence is large for society, so government should do movements for problems. Third, <b>the government can support the development</b> of technologies. <b>If government buys new technologies and uses them, more people will know about their technologies</b>. That money <b>may be used</b> for <b>the next</b> activities of <b>technology</b>. When they are common, they improve the environment. The Japanese economy will also be more active by using money. For these</p>			



それらの数字と質問紙アンケートの多肢選択の回答を量的に分析する一方で、質問紙アンケートの自由記述部分を質的に分析する方法を採用した。分析は授業に携わった教員2名が担当した。分析結果は5.1の表1に記載している。表1の「総語数」とは、各学生が3回書いたエッセイの総語数の平均を表している。「総語数の変化」とは、自力で書いた英文（修正前）と Tr もしくは WIC を使って修正した英文（修正後）の総語数の増減をパーセンテージで表したものである。「語のバリエーション」とは、総語数から重なり語数を差し引いた異なり語数のことである。「語訂正数」とは修正後に訂正した語数の全167枚（3回分）のワークシートに記入させた数の平均である。これは図3の赤で示された部分であり、図においては、make, waterなどの語を訂正している。更に、「表現訂正数」とは、二語以上を訂正した数の平均であり、図3においては緑で示された a good influence on the environment が相当する。「文訂正数」は、実践前後の文訂正数の平均である。図4では、青で示した More people assist action for nature, too. を表している。

## 5. 結果

### 5.1. Tr と WIC を使った学生の英文の比較

Tr 群と WIC 群において得られた学生の英文の質の変化は次の通りである。

表1 各ツールの平均比較（総語数の変化は割合、それ以外は語数で示す）

		Write and Improve	Transable
総語数	修正前	149	181
	修正後	167	192
総語数の変化%		10	6.5
異なり語数	修正前	85	95
	修正後	95	100
語の修正		3.9	3.9
表現の修正		2.4	4.7
文の修正		4.3	4.2

表2 各項目の変化との相関係数比較

	Write and Improve	Transable
総語数	0.58	-0.01
総語数変化	0.41	0.1
異なり語数の 変化	0.19	0.16
誤り率	0.94	-0.35
語の修正	-0.86	0.13
表現の修正	0.56	0.16
文の修正	0.52	0.16

自動採点ツールの相関の比較の計算の方法は、A: CEFRレベルをA1=1点, A2=2点, B1=3点, B3=4点, C1=5点, C2=6点, とし, 修正前と修正の後でどの程度CEFRのルーブリックの点数が上がったのかを比較して差を算出した。また, B: 各項目である総語数, 異なり語数, 語の訂正数, 表現の訂正数, 文の訂正数を修正前と修正後で比較し, 差を算出した。そして, AとBに対して相関を求めた。

表1によるとWICを使った群は, 修正後の総語数167語, 異なり語数95語, Trを使った群は総語数192語, 異なり語数100語で, Trを使った群の方が, 総語数が多く, 異なり語数も多い。しかし総語数の変化を比較すると, WIC 10%, Tr 6.5%でWICの方が総語数の変化が大きい。語の訂正数, 文の訂正数についてはほぼ違いはないが, 表現の訂正がWICを用いた群が2.4語, Trを用いた群の方が4.7語であることから, Trの方が表現の訂正が多い。つまり研究の目的1. 学生の英文の質は向上するか, に関しては, 現在商用化されている既存のWICの方が総語数の伸びに関して優れているものの, 表現の訂正に関してはTrが数を上回っていると言える。この理由としてTrにはGrammarlyが搭載されており, Grammarlyは表現の誤りについて赤線で指摘すると同時に, 訂正の候補が提示されるため修正数が伸びたと考えられる。また, Trは「表現を聞く」という箇所異なる表現を調べて学生が訂正した可能性もある。WICの場合は, 修正の指摘はあるが, 候補は提示されないところから, Grammarlyの添削機能が搭載されている点においてTrの有利な点が示唆された。

表2は表1に示した各項目の数字の変化と、TrおよびWICそれぞれに搭載されたCEFRルーブリックに基づくレベル変化の相関関係の係数を調べたものである。どの項目を訂正すれば修正前後でCEFRレベルが変化をする確率が高いかを算出したものである。表2によるとWICは総語数とCEFRレベルの上昇の相関が $r=.58$ とやや強めの相関が出ている。これは、WICが訂正後に語数を増やすと、CEFRレベルが上がりやすく、英文の語数が増えるほど良い評価が得られやすいことがわかる。また、誤り率とCEFRによるルーブリック評価の相関が $r=.94$ であることは、誤りの訂正数が多いほど、CEFRレベルが上がりやすいことがわかる。文の修正に関しても $r=.529$ が出ており、文全体を直すと評価が上がりやすい。一方、Trをみるとこれらの項目に対して相関関係があまり見られない。中には総語数にマイナスの相関も見られることから、語数を単に増やしてもレベルが下がることもあることがわかる。これについては、ChatGPTは内容面に関する評価ができることが示唆されるため、内容面を主に重視した評価がなされていることが予測される。そのため文法に関する項目との相関が起こりにくい可能性がある。ChatGPTを搭載したTrには、CEFRのほかにもGTECやCEFRの評価ルーブリックを読み込んでいる。評価ルーブリックには様々なものがあるが、総じて内容構成面における採点と、文法面における採点が組み込まれているとされている。内容面の採点のアルゴリズムを精査したわけではないが、文法面の要素との相関関係が弱いところから、Trは間接的に内容面を加味していることが考えられる。実際に、学生の提出したワークシートのうち、Trに搭載したChatGPTの評価によれば、「主題の明確さと論理的な構成があり、具体的で説得力があるが、具体的な例や統計データを挙げると信憑性が高まる」等、文法や語彙に対する指摘ではなく、パラグラフを構成する主題や、論理的な流れなど内容面に関する評価を行っていることがわかり、Trの自動採点は単に文法の訂正や語句の訂正のみを促す添削よりも、学生の英文の質を上げることにに関して有益だと言えるであろう。

## 5.2. 質問紙アンケートの結果

質問紙アンケートの項目は次の通りである。

1. 各ツールの機能のうち最も便利だった機能はどれですか。
2. 各ツールの評価機能はどうでしたか。
3. 各ツールを使用することで何が期待できますか。
4. 各ツールを今後も使い続けたいですか。

まず、最も便利な機能については文法チェックを挙げた学生がWIC 43.3%, Tr 42.3%であり、評価機能についてはWIC 92%, Tr 83%がほぼ満足していると答えた。各ツールを使うことで期待できることに関しては「英作文の質が良くなる」と答えた学生がWICは60%, Trは76.9%いた。また、「今後も使い続けたいか」という項目に関しては、「使い続けたい」とした学生は、WICが34.6%, Trが30%であり、また使いつつも英語力をあげたいという学生と合わせると両者とも6割以上をしめた。残りの学生は、使ってもいいが他のツールと併用したいと答えている。学生の自由記述による回答を分析すると、他のツールとは、電子辞書や高校時代に使っていた文法書なども挙げており、既存のものと組み合わせながら有効的に活用したいという意見も全体の2割ほど見られた。彼らによると、TrはDeepLを搭載していたが、機械翻訳に全面的に頼るのではなく、自らの英語力を向上させるために辞書や文法書も活用しながら、自分の英語力に相応する英文を作成したいと答えている。機械翻訳に頼らない英語力を望む前向きな姿勢も伺えた。また、英語が苦手な学生が、機械翻訳が足場かけとなり、英文ライティングに対する不安が軽減したと答えていることは注目できる。更には、WICにもTrにも両方搭載されている自動採点添削ツールに関して、即時に採点添削が返却されることを高く評価している一方、特にTrの評価のブレに関して懐疑心を抱いている学生もいた。しかしそのような学生も、何度も評価が上がることを期待して書き直しをしており、ライティング支援ツールが学生のモチベーションを高める一助となったことが伺える。

### 5.3. 考察

研究の目的1. Trを用いた英作文で英文の質が向上するか、の問いに関しては、Trを利用したことで全体の総語数や異なり語数が増え、Grammarlyの機能の搭載によって修正の候補が提示されている点で、表現の訂正数の伸びが見られ、修正の指摘のみで候補は提示しないWICと比較して、Tr

の有用性が確認された。また研究の目的2.自動採点は適切に評価しているか、に関しては、WICは文法面の誤りの訂正とCEFRの評価の上昇に関して比較的強めの相関が見られたが、Trにはそのような相関が見られなかったことから文法以外の側面で自動採点されていることが予測された。これについては、学生の提出したワークシートのうち、Trの中に搭載されているChatGPTが出した評価に、「主題の明確さと論理的な構成の評価」「具体的な例や統計データ要」など内容面に関する評価を行なっていることから示唆される。しかし、Trは開発段階であり評価の安定性に疑問も出た。次に、研究の目的3.学生が感じたWICとTrの使用感に関しては、質問紙アンケートの結果から、英文の質が上がり、ライティング学習に有益であり、今後も使い続けたいと答えた学生が多数いたことは着目に値する。

しかし、今回の実践実験授業に関しては、研究の限界があったことは否定できない。60名を対象とし限られた数の学生が書いた英文の一側面のみを調査しただけに過ぎず、また内容に関する項目を数式にして結果を算出したわけではない。そのためTrが内容面の採点ができているのかに関しては次なる検証が待たれる。

## 6. 結論

実践実験授業により、Trは内容面における自動採点を行っている可能性があることがわかった。また、TrもWICも学生のライティングエッセーの英文の質を向上させる有用なツールであることが伺えた。今回の実験はあくまで小規模なサンプルにおける実践実験授業であり、内容面の複雑な計算式を利用してデータを算出しておらず今後の検証が待たれる。しかし、TrやWICなどの自動添削および自動採点ツールは文法や語彙などの指摘をしてくれるため、教員のフィードバックの負担を軽減するとともに、指導をより論理性および内容面などに焦点化できる可能性が大きい。また機械翻訳は苦手意識のある学生にとって英文ライティングに対する不安を軽減し、モチベーションを高めること(Niño, 2008, 2009)もわかった。今回の実践においてWICおよびTrにおける授業での有用性が示唆され、今後のライティング活動における機械翻訳や生成系AIによる自動添削・自動採点の利用が示唆された。ChatGPTに代表される自動添削・自動採点は、内容・

構成などのライティング構成要素を添削するため、学習者自身が英文の質を向上させるのに役立つと言える (Mizumoto & Eguchi, 2023)。さらに、ライティング学習にとって大切なことは自律的な書き手を育てることであるとする、佐渡島・太田 (2013) の主張が支持され、自動添削・採点ツールの授業内での今後の利用が期待できるであろう。

## 注

1. GTEC (Global Test of English Communication): ベネッセのスコア型英語四技能テスト
2. IELTS (International English Language Testing System): 英国圏の国々に留学、就労、または移住希望の人の英語力を測定する四技能の英語試験
3. CEFR: 外国語の学習・教授・評価のためのヨーロッパ言語共通参照枠で、言語の枠や国境を越えて、外国語の運用能力を同一の基準で測ることが出来る国際基準

## 参考文献

- Burstein, J., Tetreault, J., & Madnani, N. (2013). The e-rater® automated essay scoring system. In M. D. Shermis & J. Burstein (Eds.), *Handbook of automated essay evaluation* (pp. 55-67). Routledge.
- Condon, W. (2013). Large-scale assessment, locally developed measures, and automated scoring of essays: Fishing for red herrings? *Assessing Writing*, 18(1), 100-108. <http://doi.org/10.1016/j.asw.2012.11.001>.
- 土肥康輔・須藤克仁・中村哲 (2023). 「エッセイ自動採点における文法特徴と学習者レベルの関係」『情報処理学会 第29回年次大会 発表論文集』, 211-215.
- ETS. (2021). *Criterion* [Online writing tool]. <https://etsjapan.jp/criterion/>
- Ferris, D. (2007). Preparing teachers to respond to student writing. *Journal of Second Language Writing*, 16, 165-193.
- Kilmova, B., & Pilhart, M. (2022). Neural machine translation in foreign language teaching and learning: A systematic review. *Education and Information Technologies*, 28, 663-682.
- Lee, S. M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157-175.
- Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, 35(4), 605-634.
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI lan-

- language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050. <http://doi.org.10.1016/j.rmal.2023.100050>
- Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, 21(1), 29-49.
- Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241-258. <http://doi.org.10.1017/S0958344009000172>
- 小田登志子 (2021). 「機械翻訳と共存する外国語学習活動とは」『東京経済大学 人文自然科学論集』145, 3-27.
- Page, E. B. (1967). The imminence of grading essays by computer. *Phi Delta Kappan*, 48, 238-243.
- 佐渡島沙織・太田裕子 (2013). 『文章チュータリングの理念と実践——早稲田大学ライティング・センターでの取り組み』ひつじ書房.
- Tsai, S. (2020). Chinese students' perceptions of using Google Translate as a translanguaging CALL tool in EFL writing. *Computer Assisted Language Learning*, 35(5-6), 1250-1272. <http://doi.org.10.1080/09588221.2020.1799412>
- 山田優 (2021). 「ポストエディットと持続可能な翻訳の未来」『関西大学 外国語学部紀要』24, 83-105.

(滋賀県立大学非常勤)

kenchami2@gmail.com

(立命館大学)

mihoron@fc.ritsumeai.ac.jp