

# EIKEN BULLETIN

vol. 37 2025

## 第37回「英検」研究助成 報告

### A. 研究部門

- 英語能力テストに関する研究

### B. 実践部門

- 英語能力向上をめざす教育実践

### C. 調査部門

- 英語教育関連の調査・アンケートの実施と分析

### 第37回「英検」研究助成 選考委員 (五十音順)

\*各委員の所属・役職は委嘱当時のものです。 ※は専門選考委員となります。

- ※ 和泉 伸一 上智大学 教授
- ※ 伊藤 寛和 公益財団法人 日本英語検定協会 理事・事務局長
- 岩崎 紀美子 全日本中学校長会 編集部 副部長
- ※ 小泉 利恵 筑波大学 教授
- ※ 斉田 智里 横浜国立大学 教授
- ※ 竹内 理 関西大学 教授
- ※ 寺内 一 高千穂大学 学長
- 富永 幸 文部科学省 初等中等教育局 教科調査官
- 中村 勝徳 全国英語教育研究団体連合会 会長
- ※ 西垣 知佳子 千葉大学 教授
- 藤原 和美 全国高等学校長協会 英語部会代表

---

## はじめに

---

関西大学 副学長

竹内 理

今や生成AI活用の全盛の時代である。いや、生成AI活用の全盛期が始まった、というべきかもしれない。教育の面では、音声や作文の採点・フィードバックにAIが取り入れられはじめている。また、試験問題の作成などにも、少なからずAIが役立っている。

もちろん研究の面でも、ご多分に漏れず、生成AIの活用が積極的に行われている。たとえば筆者は、分析データの下準備や分析用のコードの作成、データの可視化のための図表の作成などに、AIを大いに活用している。また聞くところによると、プレゼン用のスライドを作成させたり、データの分析をさせたり、さらには論文のたたき台を書かせたり、というような利用も当然のように行われているという。

このような中で、AI利用実態を研究者に申告させようとする動きも加速化している。2024年に筆者らの論文が受理された国際雑誌には、論文末尾にDeclaration of generative AI and AI-assisted technologies in the writing processという必須記入項目があり、AI使用の度合いを投稿者に自ら申告させている。筆者は、

The authors used Grammarly and ChatGPT to write this manuscript to improve the clarity and consistency of language usage. After using these tools, the authors reviewed and edited the content as necessary and take full responsibility for the publication's content.

という説明を書き入れて、使用実態を報告した。しかし、これと似た記載がある論文であっても、実際はAIにかなりの部分を任せているケースも存在しているようだ。もっと極端な事例では、筆者が査読を担当した国際誌への投稿論文の中に、存在しない参考文献が数件引用されているケースがあった。さらに、文体が途中から明らかに変わる事例なども、ここ1-2年の査読で急激に増えており、驚くことも多い。ただ、このあたりは、まだ人間の査読者で見破れるケースだ。もう数年もすれば、人間の目では見破られないdeepfake的な論文も横行することだろう。

さて、このような時代において、私たちはどうAIに向き合っていくのが良いのだろうか。どこまでAIにやらせるのが妥当なのだろうか。また、査読では何を見ていけば良いのだろうか。そして、私たち人間の独自性は、どこで示すべきなのだろうか。線引きと責任を巡る深刻な問題が、目の前に横たわっている。

時代は常に私たちの先を行く。これらの問題に対して、私たちは早急に、しかもAIの進歩に合わせて繰り返し議論を重ね、コンセンサスを見出していかなければならない。英検助成を受けられたAspiring Researchersの皆さんにも、自分の専門分野だけを見るのではなく、ぜひこうした議論に参加していただき、積極的に意見を述べてほしい。そんな思いを巡らせながら、この「はじめに」の責を果たすことにしたい。

## Contents

### ● はじめに

関西大学 副学長 竹内 理 …… 3

### ● 報告別講評 専門選考委員 …… 6

## A 研究部門

### 英語4技能のパフォーマンスにおける明示的知識, 自動化された明示的知識, および暗示的知識の相対的重要度の解明

研究者：石原 健志 大阪府／大阪星光学院中学校・高等学校 教諭 …… 13

### 高等学校の観点別学習状況の評価における教室内英語テストの妥当性検証と改善

研究者：大鋸 雄介 神奈川県／神奈川県立七里ガ浜高等学校 教諭 …… 33  
東京都／東京学芸大学連合大学院 在籍

### 生成 AI を用いたルーブリック評価の妥当性と可能性の探究： 教員評価と生成 AI による評価の比較研究

研究者：大和田 彩 高知県／高知県立室戸高等学校 教諭 …… 65  
(申請時：米国／ハワイ大学マノア校大学院 在籍)

### 基本的心理欲求の充足が小学校における外国語学習の楽しさと 学習者エンゲージメントに与える影響

研究者：中西 洋平 大阪府／関西大学大学院 在籍 …… 81

### 多肢選択式読解問題における本文の語を含む割合を調整した正答選択肢の調査 ー 生成AIを用いた選択肢の作成方法の検討 ー

研究者：若松 千智 千葉県／千葉大学大学院 在籍 …… 97

## Contents

### B 実践部門

#### 高等学校における CLIL を用いた思考力・表現力を養う授業の開発と評価

研究者：出尾 美由紀 英国／ウォーリック大学大学院 在籍 …… 117  
(申請時 大阪府／港南造形高等学校 教諭)

#### 教科横断型英語指導によるaudience awareness と語彙の育成

##### ー 家庭科との横断を通して ー

研究者：鳥居 敦子 大阪府／大阪教育大学附属池田中学校 教諭 …… 137

#### イントネーション指導を統合した音読活動：

##### 発話のComprehensibilityと語彙保持に及ぼす効果

研究者：根子 雄一郎 東京都／東京大学教育学部附属中等教育学校 教諭 …… 164

#### 個別最適な学びを実現する複線型の英語科授業を志向した探究的实践

研究代表者：濱田 活仁 愛媛県／弓削商船高等専門学校 助教 …… 183

共同研究者：新美 徳康 愛知県／豊田工業高等専門学校 講師  
(申請時 広島県／広島大学大学院 在籍)

福光 将仁 茨城県／筑波大学大学院 在籍

#### 高等学校のライティング指導における生成 AI の活用とその効果

##### ー 機能的適切さ(Functional Adequacy)に着目して ー

研究者：渡邊 大志 神奈川県／神奈川県立総合教育センター 指導主事 …… 207  
(申請時 神奈川県立藤沢工科高等学校 教諭／横浜国立大学教職大学院 在籍)

### C 調査部門

#### 動機づけ方略に対する英語教員の認識と実際の使用の関係についての考察

研究者：川光 大介 大阪府／大阪公立大学工業高等専門学校 講師 …… 233  
大阪府／関西大学大学院 在籍

● 今回の報告について 日本英語検定協会 名誉顧問より …… 250

## 報告別講評

### A. 研究部門・報告Ⅰ

西垣 知佳子

#### 英語4技能のパフォーマンスにおける 明示的知識、自動化された明示的知識、 および暗示的知識の相対的重要度の解明

【研究者：石原 健志】

日ごろの授業では、文法の説明はできても即興的に英語を書いたり話したりすることが難しい生徒は少なくない。その状況は文法の理解不足だけでは説明できない。本研究は、中学2年生を対象に、英語の文法知識を「明示的知識」「自動化された明示的知識」「暗示的知識」の側面から測定し、それぞれが4技能にどのように寄与するのかを構造方程式モデリング(SEM)によって実証的に検証した。その結果、特に、自動化された明示的知識が4技能すべてに安定した正の影響を示した点に注目したい。この結果は、文法規則を理解した後、繰り返しの練習と実際の使用を通して処理が速くなるというスキル習得理論の枠組みと整合しており、授業中の練習や産出活動の重要性を支持する。また、明示的知識が直接的に技能向上へ結びつくのではなく、自動化を介して間接的に効果を及ぼすという点も、授業中の文法指導と技能育成の関係を考える上で有益な示唆を与える。

一方で、暗示的知識が4技能に有意な影響を示さなかった点については、学習者の年齢・学年から考えて暗示的知識の形成が十分でない可能性や、測定方法の妥当性に関わる問題が考えられるため、今後さらなる検証が期待される。

以上のように、本研究は文法知識の質的側面を区別し、その働きを明らかにした点で高く評価される。特に、自動化の重要性を実証的に示したことは、授業での文法学習と英語運用力育成をどのように橋渡しするかを考える上で重要な手がかりを提供している。

### A. 研究部門・報告Ⅱ

小泉 利恵

#### 高等学校の観点別学習状況の 評価における教室英語テストの 妥当性検証と改善

【研究者：大鋸 雄介】

英語の定期テストやパフォーマンステストの観点別学習状況の評価は、全国で現在行われている。しかし、その妥当性はどのようなのだろうか。大鋸氏は、ある高校においてその問いを、論証に基づく妥当性検証アプローチの「領域定義推論」と「波及効果推論」の観点から詳細に分析した。その結果、妥当性の肯定的な証拠と改善点が整理された。

大鋸氏の研究の長所の一つは、教師と生徒への質問紙調査と専門家によるレビューを組み合わせ、丁寧に問いを調べたことである。内部関係者と外部の専門家の視点は異なり、それぞれを活かして解釈を深めている。この方法は、観点別学習状況の評価実践の質を高めたい小中高の教師にも参考になる。

今後は、本研究で扱えなかった論証に基づく妥当性検証アプローチの「得点化推論」、「一般化推論」、「説明推論」、「外挿推論」、「利用推論」の5点からの検討を進めてほしい。形成的評価と総括的評価を兼ねる教室英語テストには、学習進度を確認して全体を把握する役割と、学習に役立てるために生徒と教師が目標を設定し、学習後にそのプロセスと結果を振り返り、目標や学習方法を調整して次に活かすサイクルを支援する役割がある。その二つの役割のバランスをどう考え、論証立てて妥当性を確認していくか。その難題に対し、大鋸氏には、しなやかかつ大胆に取り組んでいただきたい。本研究は、その実現に向けた重要な一歩を示す卓越した成果である。

## A. 研究部門・報告Ⅲ

竹内 理

生成AIを用いたルーブリック評価の  
妥当性と可能性の探究:教員評価と  
生成AIによる評価の比較研究

【研究者:大和田 彩】

本研究は、高校生の英作文に対する教員による評価と生成AIによる評価を比較し、AIを活用したルーブリック評価の妥当性を検証することにある。その結果、文法・語彙・構成といった形式的な観点ではAIと教員の評価が高い一致度を示すが、本研究のおもしろい点は、内容や主観的態度の採点に関しては、教師の方が、AIよりも高めの評価をする傾向が示された点であろう。これは、生徒の実態を見て、良いところを積極的にくみ取ろうとする教師の行動を反映しているものと考えられ、教師が動機づけなど様々な側面を考えながら評価をしている実態を浮き彫りにしているものと思われる。一方で、形式的な側面に関しては、ルーブリックを整備することで、一貫性を維持した評価が、AIを活用することで可能であることも示した。このことは、形式面であれば、「作文回数を増やしても、AIに評価やフィードバックを任せることで、負担増を生み出さない」ということも意味しており、well-beingが重要視される時代の朗報ともいえよう。

著者は、上記の結果をしっかりとした理論的枠組みと統計手法に依拠しながら導いており、この面からも高い評価に値する。これから生成AIはますます発展していき、教師を凌駕する面が多く出てくることは間違いがないだろう。しかし、AIに任せてしまえない側面も確かに存在しているのだ。この点を明らかにし、棲み分けが必要であるということを示した本論文は、とてもタイムリーなものといえよう。

## A. 研究部門・報告Ⅳ

寺内 一

基本的心理欲求の充足が  
小学校における外国語学習の楽しさと  
学習者エンゲージメントに与える影響

【研究者:中西 洋平】

本研究は、小学校外国語教育における児童の「外国語学習の楽しさ」と「学習者エンゲージメント」に焦点を当て、その背景要因として「基本的心理欲求の充足」を位置づけた実証的研究である。

2020年度の学習指導要領改訂以降、正式教科となった外国語科の授業実態を心理的側面から捉えた意義は大きい。日本の小学校5・6年生369名を対象に、信頼性の高い既存尺度を児童向けに改訂し、部分最小二乗構造方程式モデリングを用いて分析した点も新規性が高い。結果として、基本的心理欲求の充足が「外国語学習の楽しさ」を介してエンゲージメントを高めるという媒介の関係が示され、学習の「楽しさ」が児童の主体的な学びを支える心理的基盤であることが実証的に裏づけられた。児童を対象とする複雑なモデルの検証には限界があるものの、理論と実践をつなぐ試みとして価値が高い。特に、自律性・有能感・関係性という基本的欲求が小学生段階からどのように支えられるかを探る視点は、今後の英語教育の基盤づくりに寄与するだろう。また、分析に用いられた各尺度の信頼性係数が高く、測定の厳密さが確保されている点も注目される。

今後は、授業観察や教員インタビューを通じて量的結果の背後にある要因を明らかにすることで、授業設計への具体的な示唆を提供することが期待される。児童が安心して学び、楽しさを実感できる外国語教育のあり方を探る上で、本研究は貴重な基盤を提供している。

## A. 研究部門・報告Ⅴ

斉田 智里

### 多肢選択式読解問題における本文の語を含む割合を調整した正答選択肢の調査

— 生成 AI を用いた選択肢の作成方法の検討 —

【研究者：若松 千智】

生成 AI を定期テストに活用したいと考えている教員は多いだろう。若松氏は、多肢選択式の正答肢の作成に生成 AI を活用し、本文の内容語が含まれる割合を調整することで、解答傾向がどう変化するかを詳細に検討した意欲的な作品である。生成 AI の定期テスト活用への心理的ハードルを下げることに貢献するだろう。テスト研究の観点から、次の3点を指摘しておきたい。

(1) 正答率: テストデータは、行に受検者、列に項目をとり、各セルには各受検者の各項目の解答(例: 正答=1, 誤答=0)を入力する。テストデータを横に集計すれば、受検者ごとの(得点)正答率が、縦に集計すれば、項目ごとの正答率(難易度)が示される。どちらも同じ「正答率」なので、混乱が生じる場合がある。どちらの意味で「正答率」という用語が使用されているのかを意識しながら、分析や解釈を行う必要がある

(2) 技能: 観点別評価を行う際、4技能の「技能」を、「知識・技能」と「思考・判断・表現」に分けることになる。4技能の「技能」と「知識・技能」の「技能」とはどう違うのか。教員や研究者によって解釈がまちまちといってよい。用語の表面的意味に引きずられず、各技能の構成概念を明らかにして、「知識・技能」と「思考・判断・表現」の分類根拠を明確にして作問にあたりたい。

(3) 生成 AI: プロンプトには、使用目的に対して、できるだけ細かい条件を入力し、明示的に指示をすることが求められる。あいまいなプロンプトでは、あいまいな結果しか得られない。生成 AI を活用した教室内テスト研究の一層の発展を期待したい。

## B. 実践部門・報告Ⅰ

竹内 理

### 高等学校における CLIL を用いた思考力・表現力を養う授業の開発と評価

【研究者：出尾 美由紀】

本実践は、高校において、内容言語統合型学習(CLIL)と対話型鑑賞(VTS)という枠組みを取り入れた英語授業を実践し、それが生徒の思考力と表現力の向上にどのように寄与するかを検証しようとしたものである。検証の結果、理由や根拠を伴う説明の増加、創造的・推測的発話の活発化、生徒同士の協働的な対話や構文的多様性の拡大が確認されたという。これらの変容は、文法・構文の習得にとどまらず、論理的かつ相互的な対話能力の向上にも本実践が影響している可能性を示唆しており興味深い。また、本報告では、教科横断的要素も含む実践を細やかに説明し、次に続く実践者の指針を示している点も、大いに評価に値するであろう。

本実践は、1つの(特定の)学校での試みで、しかも少人数の生徒を対象とした、対照群なしの試みであるため、状況に影響された(situatedな)結果となった点は、致し方ないと思われる。それでも、あとに続く実践への道標となるためには、今回の特定状況を詳細に記述し、類似の実践がやりやすくなるよう情報を精一杯提供することが大切である。この面においても、本論文には随所に努力の痕が垣間見られ、好感が持てる。

実践を中心にして報告するか、検証を中心にして報告するかでも、著者は大いに迷いながら、本報告を執筆したと聞いている。その葛藤の中で、両者のバランスを保った良い報告書を作り出した著者の努力に、大いに敬意を払いたい。

## B. 実践部門・報告Ⅱ

和泉 伸一

教科横断型英語指導による  
audience awarenessと語彙の育成

— 家庭科との横断を通して —

【研究者：鳥居 敦子】

英語学習指導要領に「目的・場面・状況」に注目して指導することが明記され、「目・場・状」が一言葉のように広がったことは、コミュニケーション重視の英語教育に向けた大きな前進である。言葉は本来、具体的な状況の中で有用な意味を伝え合うために使われるものであり、それを意識してこそ真のコミュニケーション能力の育成が可能になる。

その中で、ライティング指導において見落とされがちなのがaudience awareness(読み手意識)である。これは「誰のために、どんな目的で書くのか」を意識することであり、実際の手紙やメールで読み手を意識せずに書くことはあり得ないだろう。読み手が不明なままでは、何をどこまで、どのように書くかを判断できず、情報量や表現の適切さも決められない。筆者自身も論文やエッセイ執筆の際には、常に読み手を想定して内容や語り口を調整している。

鳥居氏の研究では、中学1年生を対象に、教科横断的な英語授業を通してaudience awarenessを育む指導の効果を検証した。その結果、この段階の学習者でも読み手を意識した書き方を模索できることが明らかになり、さらに、文化的背景を考慮したり語彙や表現を調整したりする兆しも見られた。

本研究は初学者を対象としたものであるが、今後は高校、さらには大学教育へこうした意識を継続的に育成していくことが重要である。こういった研究が今後も継続して行われていく中で、書き手中心ではなく、相手に伝えることを重視した「伝える・伝わる」ライティング教育が広まっていくことを大いに期待している。

## B. 実践部門・報告Ⅲ

和泉 伸一

イントネーション指導を統合した  
音読活動：発話のComprehensibilityと  
語彙保持に及ぼす効果

【研究者：根子 雄一郎】

音読活動は中学・高校の英語授業で最も多く行われている活動の一つである。しかし、その多くは「何回読むか」といった量的な側面ばかりに重点が置かれ、どのように音読すべきか、何に注意して行うべきかといった質的側面への意識は十分とは言えない。とりわけ、せっかく声に出す活動でありながら、指導が声量のみにとどまり、具体的な音声指導が行われていないことは、とても残念なことである。

こうした中、根子氏の研究は音読活動における音声面での明示的指導に焦点を当てている。イントネーションやリズム・強勢の指導を音読活動に組み込み、その効果を実験的に検証したものである。事前・事後テストとして読み上げ課題を行い、学習者の発話を「理解しやすさ」(comprehensibility)および語彙習得の観点から分析している点が特徴的である。まさに「言葉は音声である」という原点に立ち返った研究といえるだろう。

結果として、リズム・強勢・イントネーションの指導は学習者の発話の理解しやすさを高めることが確認された。また、音声指導を行う際には、学習者の英語力や教材の難易度が認知負荷に影響を与えることにも留意すべきであると指摘している。したがって、教室で音声指導を取り入れる場合は、課題がどの程度学習者の注意資源を必要とするかを見極め、段階的に指導を進めることが重要である。

音読活動を単なる反復練習として終わらせず、音声的側面に焦点を当てることで、学習者は実際のコミュニケーションでも生かせる発話能力を身につけることができるだろう。本研究が、英語教育における音声指導のさらなる発展に寄与することを期待したい。

### 個別最適な学びを実現する複線型の 英語科授業を志向した探究的実践

【研究者：濱田 活仁／新美 徳康／福光 将仁】

本研究は、英語のプロセス・ライティング活動において、「複線型」の英語授業が、他の形態の英語授業と比べてどのような違いがあるかを調査した。具体的には、学習者が学び方を選択する「複線型」授業と、指導者が学び方を指定する「単線型」授業、および「準複線型」授業を比較して、ライティングの質と方略、また、学習者の自己調整と学習方略にどのような変容をもたらしたかを調査した。参加者は高等専門学校2年生3クラス116名であった。2回の実践を行い、事前テストと事後テストを比較した。

その結果、「複線型」授業で一定程度のライティングの質の向上が確認され、「複線型」授業における自己調整による学び方の有効性が明らかにされた。また、学習者による多様な学習方略の活用と自律が、英語力習得の効果的なアプローチであることも示された。その一方で、授業でのライティング方略については、使用の拡大はほとんど見られなかった。また、「複線型」授業の一部の学習者では、特定の学び方に固執するために、ライティングの改善に遅れが見られたことから、学習者の多様な学習スタイルやニーズに応じて個別に学習方略を柔軟に適用させるための支援が不可欠であることが示されたと言える。

本研究は、従来の受動的な「単線型」授業に対して、自己調整と主体的学習を促す「複線型」授業の有用性を明確にしており、授業での応用可能性も高く、適用範囲は広範囲にわたる。個別化された英語教育の効果を示した点において、英語授業の方法に新たな展望を提供するものである。

### 高等学校のライティング指導における 生成AIの活用とその効果

— 機能的適切さ(Functional Adequacy)に着目して —

【研究者：渡邊 大志】

生成AIをどう使ったら効果的かについては、教師が集まると必ず話題になる。しかし、使っている教師でも探索的に行っている場合が多いだろう。渡邊氏は、生成AIのGoogle Geminiを用いて指導と評価を行い、生成AI活用の一つの方向性を示した。まず指導では、授業外の自学ライティング活動を設定し、生徒は自力でドラフトを書き、プロンプトを使ってAIフィードバックを受け取り、AIのアドバイスに基づいてライティングを改訂して振り返った。評価では、AIと人による採点を両方行って比較した。その結果、自学ライティング活動では、実際に取り組んだ群のみで内容点が伸び、生成AIによるフィードバックの効果が示され、生徒からは肯定的と否定的なコメントが得られた。また、生成AIは人より少し高めのスコアを付ける傾向が見られた。

渡邊氏の報告書は、指導と評価両面での効果を研究した読み応えのある作品である。記述も詳細で、今後の生成AI研究でのモデルになるだろう。従来、生成AIが苦手と考えられてきた「内容」についても、AIが有用なフィードバックを提示できる可能性があること、また、授業外の自由選択課題での学習効果は一部見られるが限定的であること等、興味深い知見を導き出した。本研究は、今後教師が生成AIを用いてどのような指導や支援、評価を行うべきかを考えるための研究の有効なステップを示した力作である。渡邊氏には今後も関連研究を続けていただきたい。

## C. 調査部門・報告 I

寺内 一

**動機づけ方略に対する  
英語教員の認識と実際の使用の  
関係についての考察**

【研究者：川光 大介】

本研究は、高等専門学校における英語教員を対象に、動機づけ方略(Motivational Strategies: MS)の効果認識と使用実態の関係、さらに使用に影響を与える要因を明らかにしたものである。

調査1では、MSの効果認識と使用頻度を比較し、両者の間に必ずしも強い相関が見られないことを示した。調査2では、教育環境や学習者特性など、既存研究で指摘された要因に加えて、高専特有の文脈的要因の存在を明らかにしており、教育現場に根ざした実践的意義が大きい。少数ながらも質問紙と自由記述を組み合わせた調査設計は精緻であり、社会的望ましきバイアスへの配慮など、研究倫理の観点からも丁寧に実施されている。Dörnyei(2001)の理論を基盤としつつ、日本の特定のEFL文脈に焦点を当てた点も評価できる。さらに、統計的分析と質的データの両方を用いて、教員の意識と行動の差を多面的に捉えた点に研究の独自性がある。調査結果をKawamitsu and Takeuchi(2024)の5因子構造と関連づけて解釈している点も、研究の一貫性を高めている。

今後は、他地域や他校種への拡張的検証を通じて、教員の信念、制度的制約、学習者の態度がどのようにMS使用に影響するのかを比較的に探ることが期待される。また、教員研修やFDプログラムへの応用を視野に入れることで、研究成果を実践に還元することも可能であろう。教育実践と理論研究との架け橋となる意欲的な研究であり、英語教育における教員研究と動機づけ理論の発展の双方に貢献することを願っている。

● 第37回「英検」研究助成は、下記の日程で行われました。

| 項 目    | 期 間              |
|--------|------------------|
| 募集期間   | 2024年2月1日～4月30日  |
| 選 考    | 2024年5月28日～7月11日 |
| 助成金贈呈式 | 2024年8月22日       |
| 研究期間   | 2024年7月～2025年8月  |
| 報告書提出  | 2025年8月          |

## 第37回 研究助成

## A 研究部門・報告 I・英語能力テストに関する研究

# 英語4技能のパフォーマンスにおける明示的知識、 自動化された明示的知識、 および暗示的知識の相対的重要度の解明

研究者: 石原 健志 大阪府／大阪星光学院中学校・高等学校 教諭

《研究助言者: 西垣 知佳子》

## 概要

本研究は、日本の中学生を対象に、文法知識の種類(明示的知識・自動化された明示的知識・暗示的知識)が英語4技能にどのように寄与するかを検討した。私立中学校2年生175名を対象に、文法性判断テストと意味性判断テストを用いて各知識を測定し、英検CSEスコアを技能の指標とした。構造方程式モデリングの分析の結果、自動化された明示的知識は、受容(聞く・読む)および産出(話す・書く)の両技能に安定して有意な正の影響を示した。一方、明示的知識は技能に直接的には影響しなかったが、自動化された明示的知識への正の影響を通じて媒介的に技能を支えていた。暗示的知識は、いずれの技能に対しても有意な寄与を示さなかった。これらの結果は、Skill Acquisition Theory が示す「明示的知識が練習によって自動化され、実際の言語使用に結び付く」という発達のプロセスを支持するものである。教育的には、文法規則を単に理解させるだけでなく、素早く正確に運用できるよう繰り返し活用する練習設計が重要であることが示唆された。

## 1 はじめに

### 1.1 英語4技能の育成と文法指導の課題

英語教育においては、聞くこと・読むこと・話すこと・書くことといった、いわゆる4技能をバランスよく育成することの重要性が強調されている。その中で、教室における英文法の扱いについて、コミュニケーション活動による暗示的指導と、説明・練習等による明示的指導とのバランスは、教師が最も悩む事項の一つである。

第二言語習得の分野では、Krashen(1982, 1985)がインプット仮説を提唱し、学習者は理解可能なインプット(comprehensible input)を十分に与えられれば、第二言語であっても無意識的に習得できると主張した。Krashenは、この無意識的な学びを「習得(acquisition)」と呼び、意識的な文法学習による「学習(learning)」と区別し、インプットこそが第二言語習得に必要な十分であるとした。これに対して、Swain(1993)は、文法形式への気づきや処理の自動化、仮説検証といった観点から、アウトプットの重要性を指摘した。

1990年代以降は、「習得」と「学習」は、それぞれ「暗示的学習」と「明示的学習」と呼ばれるようになった。暗示的学習は文法を意識せずに習得するプロセスで、その結果として直感的に使える暗示的知識が身につく。一方、明示的学習は文法形式に注目して学習する方法で、文法規則の理解を通じて明示的知識が形成される。暗示的知識と明示的知識の関係については、明示的知識が間接的に暗示的知識を助けるとする

「弱いインターフェイス」(e.g., Cintron-Valentin & N.C. Ellis, 2015)と、明示的知識が自動化され直接的に暗示的知識を助けるとする「強いインターフェイス」(e.g., Suzuki & DeKeyser, 2017)の二つの立場がある。

現在では、文法形式への焦点化(form-focused instruction: FFI)を意味中心のコミュニケーション活動と統合することが効果的であるという見解が広く共有されている(Nassaji & Fotos, 2011)。特に、文法項目の形式に注意を向けさせる明示的指導は、多くのメタ分析によりその有効性が支持されている(Goo et al., 2015; Norris & Ortega, 2000; Spada & Tomita, 2010)。もっとも、一部の研究では、明示的指導によって形成された知識が、自動化や暗示的知識の獲得につながる機会が十分に与えられない場合、時間の経過とともに減衰する可能性も指摘されている(Tode, 2007)。

これに対して、暗示的指導とは、文法規則を明示的に説明せず、学習者がインプットを通して自然に規則に気づくことで習得を促す方法である。具体的には、大量のインプットを伴う多読活動、対象となる文法項目が頻出する教材の使用、さらにはインプット活動とアウトプット活動を組み合わせることなどが挙げられる。

日本の英語教育において、明示的指導と暗示的指導のどちらをどの程度重視すべきかについては、いまだ明確な結論が得られていない。これは、「文法指導」といっても、その効果が文法項目によって異なるからである。例えば、代名詞は直前の名詞を繰り返さずに置き換えるという規則を持つが、この規則を明示的に学ばなくても、John dreamed that Bob kicked him. の him が Bob ではなく John を指すことは、多くの場合、直感的に理解できる。一方、Spada and Tomita(2010)はメタ分析により、文法項目や学習条件によって指導効果が異なり、明示的・暗示的指導の効果が一律ではないことを示した。つまり、文法項目によっては明示的知識として定着しやすいものもあれば、暗示的知識に変換されやすいものもあり、どのような知識として習得されるかは、項目の性質や測定方法に左右されることが示唆される。

英語4技能を育てるうえで、「どのような文法指導が、どの種類の知識を育て、それが技能のパフォーマンスにどう影響するのか」については、まだはっきりわかっていない点が多い。特に、文法指導で身につけた知識が、実際の言語使用の場面で使えるものになっているのか、すなわち暗示的知識として定着しているのか、あるいは明示的知識が自動的に使える状態(知識の自動化)になっているのかを十分に調べた研究は多くはない。そのため現状では、文法知識と4技能との関係を詳しく捉えるには限界がある。今後は、文法指導が4技能の育成において、どのような条件のもとで、どのような知識を通じて効果を発揮するのかを明らかにするために、知識の種類ごとに技能との関わりを調べていく必要がある。実際に、先行研究では課題実施前の明示的な文法指導が、その後の学習活動における知識の適用を促すことが報告されている(Li, Ellis, & Kim, 2018)。本研究でも、英語の関係代名詞を未学習の学習者に対し、課題前に明示的な指導を行い、その後の練習活動を通じて知識の自動化を促すことを試みた。

## 1.2 明示的知識、自動化された明示的知識と暗示的知識の違い

本節では、文法の明示的知識、自動化された明示的知識、暗示的知識の違いを、それらの習得経路、処理の意識レベル、測定可能性から整理する。まず明示的知識は、指導者による説明や教科書の解説などを通じて、意識的・メタ言語的に習得される(Ellis, 2004)。これは宣言的知識に相当し、学習者が自覚的に保持し、ルールとして説明可能な点が特徴である(DeKeyser, 2003; Ellis, 2005)。たとえば「主語が三人称単数のとき、現在形の動詞には-(e)sが付く」といった規則のように、文の組み立てや語形変化等の規則を説明できる知識を指す。つまり、明示的知識とは、文法的に適切かどうかを判断する根拠となる知識である。このような明示的知識の性質を踏まえた測定方法として、文法的に正しい文と誤った文を見分ける文法性判断テストがある(Godfroid et al., 2015; Ellis & Roever, 2018)。それ以外にも、「ある文の誤りの理由」を説明させるメタ言語知識テスト(Metalinguistic Knowledge Test; Bowles, 2011; Ellis, 2004)、与えられた文から文法的誤りを見つけて訂正させる誤り訂正タスク(Error Correction Task; Ellis, 2005)なども明示的知識を測定するものである。

これに対して、自動化された明示的知識は、明示的に学習された文法知識が、練習を通じて意識せずに運用できるレベルまで処理が速くなった状態を指す(DeKeyser, 2007; Suzuki & DeKeyser, 2017)。Skill Acquisition Theoryにおいては、明示的知識(宣言的知識)が、反復練習によって手続き的知識へと変換されるプロセスが想定されており、自動化された明示的知識は、その成果として位置づけられる(DeKeyser, 2007)。学習者が実際にこの知識を使って言語の理解や産出を行う際には、その都度、文法知識を意識せずとも、素早く正確に使うことができる。また、自動化された明示的知識は、条件によっては再び意識化・言語化できるという特徴を持つ(DeKeyser, 2007; Suzuki & DeKeyser, 2017)。さらに、教育的な観点から見ると、明示的知識の自動化は教室での言語活動や練習によって促すことができるため、後に述べる暗示的知識に比べて指導可能性が高いという利点がある(Suzuki, 2023)。

自動化された明示的知識の測定には、制限時間付き文法性判断課題のほか、「特定の単語が提示されたときに即座に反応する」という単語モニタリング課題(word monitoring task)や、「文を読むときに、一語ごとに二つの選択肢が与えられ、正しい方を選んで進む」という迷路課題(maze task)が用いられる。これらに共通するのは、意味内容ではなく文法形式に注意を向けさせる設計であり、学習者が知識を意識的に検索するのではなく迅速に処理できるか(処理の自動性)を測定する点にある。

一方、暗示的知識は、明示的に学習されたものではなく、理解可能なインプットの中から自然に抽出され、無意識のうちに獲得される知識である(Ellis, 2005)。暗示的知識の特徴は、学習者が自分で意識して使うものではなく、その知識を言語化することも困難である点である。また、習得までに比較的長い時間を要し、大量かつ意味のあるインプットの蓄積が前提とされている(和泉, 2016)。Ellis(2005)は、暗示的知識は無意識のうちに構築され、直感的な文法判断や自然な言語運用を支えると指摘している。また、暗示的知識の指標としては、文の意味性判断テストや自己ペース読み課題などが用いられてきた。これらのテストでは、意味が自然かどうかを判断させるため、学習者の注意は文法形式ではなく意味に向けられる点で、暗示的知識の測定に適している(Ellis, 2005)。

### 1.3 Skill Acquisition Theoryにおける文法知識の習得

本研究では、英語の文法知識が4技能(聞く・読む・話す・書く)のパフォーマンスにどのように関与するのかを検討するにあたり、Skill Acquisition Theoryを理論的枠組みとして採用する。特に本研究は、明示的知識が繰り返しの練習によって自動化され、実際の言語運用において迅速かつ無意識的に使えるようになるという、Skill Acquisition Theoryにおける知識発達のプロセスに注目する。

この理論では、学習は大きく三つの段階を経るとされる。まず、文法規則などを意識的に理解する段階(明示的知識の獲得)があり、その後、定型的な練習を通じて知識を動作として使える形に変換する段階がある。これを「手続き化」と呼び、知識を理解するだけでなく、適切な場面で即座に適用できるようにする過程を指す。最後に、その運用が繰り返されることで、処理にほとんど意識を必要としない「自動化」に至り、知識は流暢かつ正確な言語使用を支える基盤となる。

さらに、こうして形成された知識は、暗示的知識とは異なる「自動化された明示的知識」であるという報告もなされている(Suzuki & DeKeyser, 2017)。したがって、本研究では、文法知識の発達を単なる「明示的知識 vs 暗示的知識」という二項対立で捉えるのではなく、明示的指導と練習を通じて形成される中間的段階としての「自動化された明示的知識」の存在を仮定する。

Skill Acquisition Theoryはもともと認知心理学において提唱された理論であるが、第二言語習得にも応用されており、特に文法知識がどのように言語技能の実際の使用へと結び付いていくのかを説明する理論として有効である(DeKeyser, 2001, 2007; Suzuki, 2023)。本研究においても、知識の種類とその運用可能性とが、英語4技能にどのように関与するのかを明らかにするうえで、Skill Acquisition Theoryは適切な理論的基盤を提供する。

## 1.4 文法知識の種類と言語運用能力の関係

第二言語習得研究では、明示的知識と暗示的知識の役割について多くの議論がなされてきた。Ellis (2005, 2006) は、両者が異なる形で言語パフォーマンスに関与し、技能によって有効性が異なる可能性を示唆している。とりわけ、話す・聞くといった瞬時の処理が求められる技能では、意識的に文法規則を参照することが難しいため、暗示的知識や自動化された明示的知識が大きな役割を果たす。これに対して、読む・書くといった時間的余裕のある技能では、意識的に文法規則を思い出して活用できるため、明示的知識の重要性が高いと考えられている。

このことを踏まえて、明示的知識と暗示的知識が言語運用能力にどのように関与するかを検証した先行研究を取り上げたい。Zhang (2012) は、中国人大学生190名を対象に、語彙知識(広がりと深さ)および文法知識(明示的知識・暗示的知識)が英語読解に及ぼす影響を、構造方程式モデリング(SEM)によって分析した。文法知識の測定には、明示的知識を捉える文法エラー訂正課題と、暗示的知識を捉える制限時間付き文法性判断課題が用いられた。読解力は、共参照推論(代名詞や指示語が何を指すかを判断する能力)、文脈の推論(文脈から明示されていない意味を推測する能力)、要旨把握(文章全体の主旨を把握する能力)の3側面から評価された。その結果、語彙知識が読解力に強い影響を与える一方で、文法知識の寄与は相対的に小さかった。ただし文法知識の中では、暗示的知識が明示的知識よりも読解力との関連が強く、語彙の影響を統制した後も読解を支える要因となっていた。Zhang は、リーディング場面においては、文法規則を意識的に適用する明示的知識よりも、迅速かつ無意識的に処理される暗示的知識が理解に貢献する可能性が高いと結論づけている。

Elder and Ellis (2009) は、明示的知識と暗示的知識のいずれもがIELTSの熟達度と中程度の相関を持つことを示した。ただし、両者が同一の文法項目に対して同時に関連することは少なく、比較表現のような文法では暗示的知識が、関係詞のような文法では明示的知識が、それぞれIELTSスコアとの関連を示していた。このように、知識の種類ごとに関連する言語構造が異なることは、文法知識の影響が単純ではないことを示唆している。こうした傾向は Han and Ellis (1998) の報告とも一致しており、暗示的知識はSLEP(リスニング・リーディングのテスト)と、明示的知識はTOEFL(産出能力を含む)との相関が高かった。Suzuki and Sunada (2019) は、英語の関係節(主格と目的格)の正確性と処理速度について、焦点練習の形式(インプット/アウトプット)とスケジュール(ブロック練習 / 間隔練習 / ブロック練習と間隔練習のハイブリッド)が第二言語の受容能力と産出能力にどのように影響するかを検討した。その結果、インプット練習は受容能力に、アウトプット練習は産出能力に特に効果的であることが示された。これは、文法知識が繰り返しの練習を通じて手続き化・自動化されるというSkill Acquisition Theoryの枠組みと一致しており、文法知識の使用が受容能力と産出能力に貢献する過程を示している。

ここで、明示的知識と暗示的知識の中間体と位置づけられる自動化された明示的知識が、英語運用能力の向上に果たす役割について触れておきたい。まず産出能力に関しては、Suzuki and DeKeyser (2017) が、文法知識の自動化がスピーキングにおける正確性と流暢さの向上に寄与することを実証している。さらに受容能力に関しても、自動化の重要性が指摘されており、Suzuki (2023) は、文法知識への迅速なアクセスがリスニングやリーディングにおける効率的な処理を支える可能性を論じている。これらの知見から、自動化された明示的知識は受容・産出の両技能に関与し得ることが示唆される。また文法知識に限らず、Uchihara et al. (2024) は、音韻語彙知識(語の音声形式と意味を正確に結び付けて認識する能力)の自動化がリスニング理解に有意な影響を及ぼすことを示しており、自動化された知識が受容能力に広く関わることを裏付けている。

このように、明示的知識、自動化された明示的知識、暗示的知識はいずれも、言語のパフォーマンスや熟達度に寄与する可能性があるが、それぞれの役割や影響のあり方は異なることが示唆される。明示的知識は、学習初期において文法項目を理解・操作するうえで重要であり、適切な練習を通じて自動化されることで、自動化された明示的知識として実際の言語運用においてより迅速かつ安定した処理を支えるよ

うになる。一方で、暗示的知識は明示的指導や意識的な学習を経ずに形成され、文法的正確性や流暢さに無意識的に貢献する知識であると考えられている。しかしながら、これらの知識が4技能から構成される受容能力や産出能力にどのように影響を与えているかという研究は少ないのが現状である。

本研究では、ここまで述べてきた理論的・実践的背景を踏まえ、英語の関係代名詞に関する知識を導入していない学習者である中学生を対象に、明示的指導を施したうえで、その知識の自動化を促す練習を実施する。これにより、明示的知識がどの程度自動化され、さらに暗示的知識の形成や英語4技能の向上にどの程度寄与するのかを検証する。関係代名詞という文法項目を選定した理由は、中学校英語カリキュラムの中では比較的後半に導入されるため事前知識の影響を受けにくく、指導および練習による効果をより明確に捉えられると考えられるからである。また、授業進度の都合からも、この単元を扱う時期と本研究の実施時期が一致しており、実務的にも適合していた。

本研究で取り組むリサーチクエスチョン(RQ)は以下の通りである。

RQ1 明示的知識が日本人中学生の英語運用能力にどの程度貢献するか？

RQ2 自動化された明示的知識が日本人中学生の英語運用能力にどの程度貢献するか？

RQ3 暗示的知識が日本人中学生の英語運用能力にどの程度貢献するか？

## 2 方法

### 2.1 協力者

本研究には、私立中学校に在籍する2年生175名(全員男子)が参加した。参加者のうち、3級取得者が34名、準2級取得者が52名、2級取得者が2名であった。また、39名は受験級に合格してはいなかったが、話すことを除く3技能の英検CSEスコア(後述)の最小値および最大値から、3級から準2級程度と見なせる生徒たちであった( $M = 1313.13$ ,  $SD = 155.60$ ,  $Min = 761$ ,  $Max = 1684$ )。年齢は13歳から14歳であり、英語学習歴は小学校での義務教育3年生から6年生での4年間および中学校での1.5年間であった。全員が日本国内で育ち、英語以外の外国語の学習経験はなく、海外滞在歴もなかった。また、教育上特別な支援を要する生徒はいなかった。

本研究は、学校および参加者の保護者からの承認を得たうえで実施された。参加者には、ディスプレイ上に研究の説明を提示し、同意する場合は「はい」を選択して先に進む形式で、電子的同意(informed consent)を取得した。

### 2.2 マテリアル

本研究では、英文法の明示的知識、自動化された明示的知識、暗示的知識を測定するため、2.2.1節および2.2.2節で述べる2種類のテストを作成した。各テストは、生徒が教室内で筆者の監督のもと、オンライン実験プラットフォームであるCognition(<https://www.cognition.run/>)とiPadを用いて実施した。

テスト項目としては関係代名詞を扱った。この知識は、明示的な指導と反復練習による手続き化・自動化を目指して、図1-1、図1-2に示す教材およびタスクにより導入された。

## Grammar Practice 関係代名詞 who / which の学習

Basic Structures: 基本英文

I have an uncle who lives in Canada.

This is a letter which was written by Jim.

This is the book which I bought yesterday.

I spoke to the man who she did not know.

関係代名詞の作り方 <1つ目の文の先行詞に2つ目の文を続けて1つの文を作る>

**Pattern 1** 関係代名詞 主格 先行詞が「人」のとき

① I have an uncle. ② He lives in Canada.

<先行詞>

=> who lives in Canada.

<先行詞が人>の時、関係代名詞は who をえらぼう!

an uncle の後に②を続けるために he を who(関係代名詞)に変えて、繋ぐことができる

答え⇒ I have an uncle [who lives in Canada].

★ I have an uncle [who he lives in Canada].

**Pattern 2** 関係代名詞 主格 先行詞が「もの・こと」のとき

① This is a letter ② It was written by Jim.

<先行詞>

=> which was written by Jim.

<先行詞がもの・こと>の時、関係代名詞は which をえらぼう!

a letter の後に②を続けるために it を which(関係代名詞)に変えて、繋ぐことができる

答え⇒ This is a letter [which was written by Jim].

★ This is a letter [which it was written by Jim].

**Pattern 3** 関係代名詞 目的格 先行詞が「もの・こと」のとき

① This is the book. ② I bought it yesterday.

<先行詞>

=> I bought it yesterday.

which

=> which I bought △ yesterday.

<先行詞がもの・こと>の時、関係代名詞は which をえらぼう!

the book の後に②を続けるために it を which(関係代名詞)に変えて、主語の前に移動させることで、先行詞と繋ぐことができる

答え⇒ This is the book [which I bought △ yesterday].

★ This is the book [which I bought it yesterday].

**Pattern 4** 関係代名詞 目的格 先行詞が「人」のとき

① I spoke to the man. ② She did not know him.

<先行詞>

=> who she did not know him.

<先行詞が人>の時、関係代名詞は who をえらぼう!

a man の後に②を続けるために him を who(関係代名詞)に変えて、繋ぐことができる

答え⇒ I spoke to the man [who she did not know △].

★ I spoke to the man [who she did not know him].

■ 図1-1: 関係代名詞(主格・目的格)の練習シート

## Interleaved Practice

## 関係代名詞の選択①

次の英文の( )内に適切な関係代名詞を入れなさい。(ただし that は除く)

- 1) I have a brother ( ) loves to dance.
- 2) Televisions are devices ( ) show us many TV programs.
- 3) That is the building ( ) you can see over there.
- 4) He is the old man ( ) I met last week.

## 並べ替え①

次の日本語に合うように、( )内の語句を並べ替え、英文を完成させなさい。ただし、文頭に来る語(句)も小文字になっている。

1. 彼は私に話をしてくれた人です。

( is / the person / told / who / me / the story / he )

2. 彼は間違いでいっぱいレポートを提出した。

He handed in ( of / was / which / full / mistakes / a report ).

3. 鈴木先生は、皆が尊敬する先生だ。

Ms. Suzuki ( is / respects / who / everyone / the teacher ).

4. 私が登った山は美しかった。

( climbed / I / which / the mountain / was )

## 和訳①

以下の英文を日本語に直しなさい。

1. This is the book which you wanted to buy at the bookstore.

2. You can tell us about a news story which is famous in Japan.

3. I know a person who earns a lot of money in a year.

4. They are the people who I really respect.

■ 図1-2: 関係代名詞(主格・目的格)の練習シート

明示的な指導として、授業冒頭で、関係代名詞の基本的な機能(主格／目的格の区別、有生／無生の先行詞との結び付きなど)について、板書と例文を用いた解説を行った(図1-1)。続く練習では、①文中の空所に適切な関係代名詞(who, which, that)を補完させることで、言語形式の想起を促す穴埋め練習(図1-2 関係代名詞の選択①)、②与えられた日本語文の意味に対応する英文を完成させることで、明示的な文法知識の手続き化と構文処理スキルの育成を図る並べ替え練習(図1-2 並べ替え①)、および③関係代名詞を含む英文の意味を解釈させることで、言語形式と意味のマッチングを図る和訳練習(図1-2 和訳①)を行った。これらの練習はすべて、Skill Acquisition Theory(DeKeyser, 2007)および想起練習(Roediger & Karpicke, 2006)の枠組みに基づき、明示的知識の反復活性化と処理の迅速化を意図している。一連の指導は、50分授業×4時限(計200分)にわたって行われた。

## 2.2.1 英文の意味性判断テスト

暗示的知識の測定には意味性判断テストを用いた。このテストは、言語形式よりも意味解釈に生徒の注意を向けられるため、意識的な文法処理を求めないという特徴がある(Lim & Godfroid, 2015)。テストでは、表1に例示するような英文がiPadのモニター中央に提示され、生徒はできるだけ素早くかつ正確に、その英文の意味が自然か否かを直感的に判断するよう求められた。

■表1: 意味性判断テストで使用した英文の例

| 関係代名詞の格 | 先行詞 | 例文                                                                                                             |
|---------|-----|----------------------------------------------------------------------------------------------------------------|
| (1) 主格  | 有生  | He is the boy who can speak English well.<br>He is the boy who can <i>push</i> English well.                   |
| (2) 目的格 | 有生  | You are the teacher who we like very much.<br>You are the teacher who we <i>climb</i> very much.               |
| (3) 主格  | 無生  | This is the building which has a tea room.<br>This is the building which <i>cries</i> a tea room.              |
| (4) 目的格 | 無生  | Our flag is the one which you can see from here.<br>Our flag is the one which you can <i>answer</i> from here. |

注. 太字は意味的に不自然となる動詞を表す。

意味性判断テストは全80問で構成され、すべての英文に関係代名詞が含まれていた。出題は表1に示す四つのカテゴリに分類され、それぞれ20問ずつ含まれていた: (1) 関係代名詞の主格で先行詞が有生(主に人間や動物など、意思や行動主体と見なされるもの)のもの、(2) 関係代名詞の目的格で先行詞が有生のもの、(3) 関係代名詞の主格で先行詞が無生(物質・道具・抽象概念など、意志や感情を持たないもの)のもの、(4) 関係代名詞の目的格で先行詞が無生のもの。各カテゴリには、意味的に自然な文(正文)と不自然な文(非文)がそれぞれ10問ずつ含まれていた。

正文は、関係代名詞を含む英文として意味が自然に通るものであり、たとえば表1に示されている「関係代名詞主格・先行詞有生」の例文 He is the boy who can speak English well. (彼が英語を上手に話すことができる少年です)のような文が該当する。

一方、非文は、正文と同じ構文を保ったまま、意味的に不自然になるように動詞のみを変更した文である。たとえば、上記の文の動詞 speak(話す)を push(押す)に置き換えた He is the boy who can push English well. (彼が英語を上手に押すことができる少年です)は、構文的には正しいが、意味的には成立しない非文となる。この不自然さは、「押す」という動詞が通常、「ボタン」や「ドア」など物理的に押せる対象に用いられるため、「英語」のような抽象的な名詞を目的語にすると意味が通らなくなる点に起因する。

なお、本テストではすべての文が80問のセットに組み込まれていたが、提示順や意味的なバイアスを制

御するために、参加者175名をAグループとBグループに分けてカウンターバランスをとった。具体的には、Aグループで正文として提示された文は、Bグループでは対応する非文として提示され、逆にAグループで非文として提示された文は、Bグループでは対応する正文として提示された。たとえば、Aグループが He is the boy who can speak English well. を正文として提示された場合、Bグループは He is the boy who can push English well. を非文として提示される。その逆に、Aグループが「push版」を非文として提示された場合、Bグループは「speak版」を正文として提示される。これにより、すべての文が全体として等しい頻度で提示され、文の内容や意味性に起因する効果が特定の条件に偏るのを防いでいる。

各参加者の判断の記録はiPadに接続されたキーボードを使い、人差し指を「k」(正しい)、中指を「l」(正しくない)に載せて、どちらかを素早く押すよう指示した。英文は一度だけ提示され、前の英文に戻ることはできないようにした。

## 2.2.2 英文の文法性判断テスト

明示的知識および自動化された明示的知識の測定には文法性判断テストを用いた。このテストでは、意味的判断よりも文法性の即時判断に注意を向けさせるため、意識的な文法処理が求められる (Ellis, 2005)。表2に例示するような英文がiPadのモニター中央に提示され、生徒はできるだけ素早くかつ正確に、その英文が文法的に正しいか否かを素早く判断するよう求められた。表2に文法性判断テストで使用した英文の例を示す。

■表2: 文法性判断テストで使用した英文の例

| 代名詞の格  | 先行詞 | 例文                                                                                               |
|--------|-----|--------------------------------------------------------------------------------------------------|
| (1)主格  | 有生  | Kazu is the boy who likes music.<br>Kazu is the boy who <b>he</b> likes music.                   |
| (2)目的格 | 有生  | He is the boy who I saw yesterday.<br>He is the boy who I saw <b>him</b> yesterday.              |
| (3)主格  | 無生  | This is a book which has sold well.<br>This is a book which <b>it</b> has sold well.             |
| (4)目的格 | 無生  | Tom found the key which I lost yesterday.<br>Tom found the key which I lost <b>it</b> yesterday. |

注. 太字は文法的に非適格となる代名詞を表す。

文法性判断テストも意味性判断テストと同じように全80問、すべての英文に関係代名詞が含まれていた。出題は表1に示す四つのカテゴリに分類され、それぞれ20問ずつ含まれていた：(1)関係代名詞の主格で先行詞が有生(主に人間や動物など、意思や行動主体と見なされるもの)のもの、(2)関係代名詞の目的格で先行詞が有生のもの、(3)関係代名詞の主格で先行詞が無生(物質・道具・抽象概念など、意志や感情を持たないもの)のもの、(4)関係代名詞の目的格で先行詞が無生のもの。各カテゴリには、文法的に適格な文(正文)と非適格な文(非文)がそれぞれ10問ずつ含まれていた。

非文の作成にあたっては、Gass(1982)に従い、関係節内に代名詞が残存する構文違反の形を採用した(表2参照)。このような文は、関係代名詞が先行詞の役割を担っているにもかかわらず、関係節内にその指示対象を繰り返す代名詞が出現するため、構文的に許容されないものとされる。

文法性判断テストでも同様に、参加者をAグループとBグループに分けてカウンターバランスをとった。具体的には、Aグループで正文として提示された文は、Bグループでは対応する非文として提示され、逆にAグループで非文として提示された文は、Bグループでは対応する正文として提示された。たとえば、Aグループが He is the boy who I saw yesterday. を正文として提示された場合、Bグループは He is the boy who I saw him yesterday. を非文として提示される。その逆に、Aグループが「非文」を提示された場合、Bグループ

ブは対応する「正文」を提示される。これにより、各文が全体として等しく提示されるよう統制され、構文的文法性の効果がグループ間で偏らないように設計されている。また、意味性判断テストと同じく、各英文は一度だけ提示され、前の英文に戻ることはできないようにした。

### 2.2.3 英語4技能スコア

参加者の英語4技能(聞くこと・読むこと・話すこと・書くこと)の評価には、英検CSEスコアを用いた。「話すこと」のスコアがない場合には、リストワイズ削除(当該参加者を分析対象から除く処理)は行わず、欠測値として記録した。CSEスコアは公益財団法人 日本英語検定協会によって提供された標準化スコアであり、項目応答理論に基づいて技能横断的な比較や合算が可能である(文部科学省, 2015; 日本英語検定協会, n.d.)。各技能は1,000点満点でスコア化されており、4技能の合計は最大4,000点となる。

## 2.3 手順

本研究のデータ収集は2024年1月に行った。生徒はまず意味性判断テストに解答し、その後、文法性判断テストに解答した。この順序にした理由は、文法性判断テストを先に提示すると文法への注意が過剰に向き、意味性判断テストにおける無意識的な文処理に影響を及ぼす可能性があったからである(Bowles, 2011)。

テストは、筆者が勤務する中学校の4クラスにおいて、1日2クラスずつ、短縮授業の期間(2024年1月16日、17日)を使って2日間にわたり行った。静かな通常教室において筆者が試験監督として同席し、全生徒に共通の指示を口頭で与えた。特に、テストの解答は「素早く行う」ことをテスト画面上でも指示し、筆者からも「できる限り素早く判断するように」と口頭で強調することで、時間制限のある課題であるという意識づけを行った。ただし、関係代名詞が生徒にとって新規文法項目であったため、制限時間以内に解答しないと不正解となるような設定は行わず、無解答が出ることを防ぐ配慮をした。

各テストの所要時間はおおよそ20分であった。それぞれのテストの開始前には、練習問題を1問、クラス全体で一緒に行い、筆者の指導のもと、解答操作と判断方法を確認した。項目順序は生徒ごとにランダム化されており、反応の順序効果を排除した。また、意味性・文法性の成立状況に関する刺激文セットは、2クラスずつで内容を反転(カウンターバランス化)しており、構成上のバイアスを最小限に抑えた。

## 2.4 分析

### 2.4.1 採点およびデータの前処理

本研究では、意味性判断テストおよび文法性判断テストの結果から、構造方程式モデリングに用いる各指標を次の手順で記録した。まず、意味性判断テストおよび文法性判断テストの正誤は0/1で採点した。次に、「正答かつ妥当な反応時間」を得るために、Godfroid et al.(2018)の基準に従い、不正解の解答、および早すぎる(500ミリ秒未満)または遅すぎる(50,000ミリ秒以上)解答を除外した。

前処理後のデータを用いて、次に述べる手順で明示的知識、自動化された明示的知識、暗示的知識の測定値を算出した。まず明示的知識については、文法性判断テストにおける非文に対する正答率(Correct Rejection率)を算出した。これは、誤りのある文を明示的に識別する能力が文法規則の理解と意識的な適用を反映するためである(Ellis, 2005)。

自動化された明示的知識については、文法性判断テストにおける正文に対する正答率(Hit率)を算出した。これは、素早い文処理が求められる状況でも文法的な英文を正しく判断できる、すなわち、素早くアクセスできる知識を有していると解釈できるためである(Suzuki & DeKeyser, 2017)。

暗示的知識は、意味性判断課題において、意味的に自然な文と不自然な文それぞれの平均反応時間(RT)を算出し、その差分値(不自然文RT - 自然文RT)を指標とした。暗示的知識が高い場合、不自然な文では解釈に困難が生じるため、自然な文よりもRTが長くなる傾向がある(Lim & Godfroid, 2015)。また、

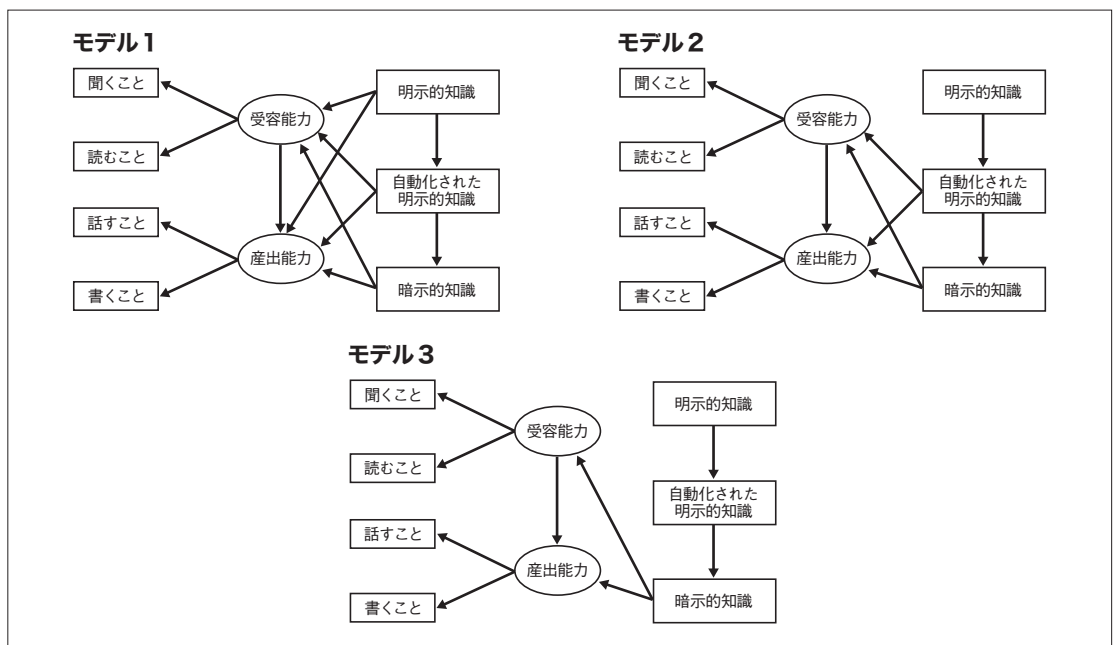
Suzuki (2017) は、暗示的知識と自動化された明示的知識を区別するためには、課題の性質や RT の分析方法が重要であることを指摘しており、本研究もこの理論的枠組みに基づいて暗示的知識を測定した。

暗示的知識の測定には、反応時間の変動係数を指標とする方法も広く用いられている。変動係数は、各参加者の反応時間の標準偏差を平均値で割った値であり、処理のばらつきを示す指標である。これは、処理の安定性を反映し、自動化の進展を測定できるとされる (Hulstijn, 2002; Hulstijn, Van Gelderen, & Schoonen, 2009)。しかし、Hui (2019) は、学習初期の変動係数がテスト開始直後に低く、その後一時的に上昇し、終盤で再び低下するという逆 U 字型の推移を示すことを報告している。これは、初期には既存知識で素早く反応し、その後新たに学んだ知識を試行錯誤的に適用し、最終的に運用が安定するという学習過程を反映していると考えられる。そのため、単純にテスト全体の平均変動係数を算出するだけでは学習過程を正確に捉えられない可能性がある。以上の理由から、本研究では変動係数ではなく、反応時間の差分を指標として用いた。

## 2.4.2 構造方程式モデリングによる分析

RQ に対する分析として構造方程式モデリングを採用した (R-4.5.1 の lavaan パッケージを使用)。まず、英語運用能力の構成概念として、受容能力 (聞くこと・読むこと) および産出能力 (話すこと・書くこと) を潜在変数とした。そして、明示的知識、自動化された明示的知識、暗示的知識が、それぞれの能力に貢献すると仮定したモデルを立てた。具体的には、図2に示す通り三つの構造モデルを立てた。モデル1では、3種の知識指標がそれぞれ受容能力および産出能力に直接影響を与え、モデル2では、明示的知識が自動化された明示的知識を経由して各能力に影響を及ぼす媒介モデル、モデル3では、さらに暗示的知識を介在させる3段階モデルとした。テスト間でスケール (満点の値) が異なるため、すべての観測変数を標準化 (z 値変換) して解析に用いた。

モデルの適合度指標には CFI, TLI, RMSEA, SRMR を算出し、その評価には Hu and Bentler (1999) が提案するカットオフ基準 ( $CFI/TLI \geq .95$ ,  $RMSEA \leq .06$ ,  $SRMR \leq .08$ ) を参照した。CFI と TLI は .95 以上で良好な適合、RMSEA と SRMR は .08 未満で誤差の少ないモデルと判断される (小林・濱田・水本, 2020)。さらに、モデルの比較には尤度比検定および情報量基準 (AIC・BIC) を併用し、統計的かつ理論的に妥当なモデルを選択した。



■ 図2: 本研究で比較した三つの構造モデル

### 3 結果

#### 3.1 記述統計と相関

各変数の記述統計量を表3に示す。英検CSEスコアにおいては、「読むこと」の平均値は462.59 ( $SD = 41.97$ )であり、「聞くこと」および「書くこと」も同程度の水準を示した。一方、「話すこと」のスコアでは歪度が1.05とやや大きく、分布に右方向への偏りが見られた。また、暗示的知識は外れ値の影響により標準偏差(1204.17)および尖度が極めて大きく、分布が非正規である可能性が示唆された。明示的知識および自動化された明示的知識はおおむね正規に近い分布を示していた。

■表3: 各変数の記述統計量

| 変数          | <i>n</i> | 平均     | <i>SD</i> | 中央値    | 最小値      | 最大値     | 歪度    | 尖度    |
|-------------|----------|--------|-----------|--------|----------|---------|-------|-------|
| 読むこと        | 131      | 462.59 | 41.97     | 465.00 | 354.00   | 565.00  | 0.07  | -0.34 |
| 聞くこと        | 131      | 432.67 | 59.80     | 432.00 | 324.00   | 591.00  | 0.37  | -0.17 |
| 書くこと        | 131      | 449.89 | 55.45     | 456.00 | 321.00   | 600.00  | 0.23  | 0.05  |
| 話すこと        | 113      | 412.01 | 44.49     | 407.00 | 340.00   | 550.00  | 1.05  | 1.15  |
| 暗示的知識       | 173      | 414.90 | 1204.17   | 342.99 | -3077.44 | 6605.97 | 0.71  | 3.80  |
| 自動化された明示的知識 | 175      | 0.73   | 0.14      | 0.74   | 0.10     | 0.98    | -0.66 | 1.17  |
| 明示的知識       | 175      | 0.42   | 0.22      | 0.40   | 0.00     | 0.98    | 0.47  | -0.54 |

次に、4技能および知識指標間の相関を表4に示す。4技能間には中程度から強い正の相関が見られ、特に「読むこと」と「聞くこと」( $r = .70$ )、「読むこと」と「話すこと」( $r = .69$ )の相関が強かった。知識指標との関係では、自動化された明示的知識は「読むこと」( $r = .52$ )、「聞くこと」( $r = .43$ )と比較的強い相関を示し、暗示的知識はすべての技能と弱い相関しか示さなかった。また、明示的知識は、4技能との相関がいずれも弱く(読むこと:  $r = .14$ , 聞くこと:  $r = .07$ , 書くこと:  $r = .04$ , 話すこと:  $r = .05$ )、特定の技能との強い関連は認められなかった。また、明示的知識と自動化された明示的知識の間には有意な相関( $r = .21$ )が見られており、明示的知識が自動化された明示的知識の基盤となっていることが示唆された。

■表4: 4技能および文法知識指標間のピアソンの相関係数

| 変数            | 1.読むこと | 2.聞くこと | 3.書くこと | 4.話すこと | 5.暗示的知識 | 6.自動化された明示的知識 | 7.明示的知識 |
|---------------|--------|--------|--------|--------|---------|---------------|---------|
| 1.読むこと        | -      | .70    | .44    | .69    | -.13    | .52           | .14     |
| 2.聞くこと        |        | -      | .40    | .64    | -.20    | .43           | .07     |
| 3.書くこと        |        |        | -      | .34    | -.21    | .05           | .04     |
| 4.話すこと        |        |        |        | -      | -.08    | .36           | .05     |
| 5.暗示的知識       |        |        |        |        | -       | -.01          | .04     |
| 6.自動化された明示的知識 |        |        |        |        |         | -             | .21     |
| 7.明示的知識       |        |        |        |        |         |               | -       |

注. 太字は統計的に有意( $p < .05$ )な相関係数を表す。

### 3.2 モデルの比較

明示的知識、自動化された明示的知識、暗示的知識が第二言語技能に与える影響を検討するため、三つのモデルを比較した。表5に示す通り、モデル3はモデル1やモデル2と比べて適合度指標の値は良くないことがわかった。Satorra-Bentlerによる尤度比検定の結果においても、モデル3はモデル1( $\chi^2[8] = 6.87, p < .001$ )およびモデル2( $\chi^2[10] = 7.47, p < .001$ )と比較して、有意に適合度が劣っていた。モデル1とモデル2の適合度は統計的に同等であったが( $\chi^2[2] = 0.60, p = .741$ )、モデル2はより少ないパラメータで同等の適合度を達成していたため、最も儉約的で妥当なモデルであると判断される。したがって、以降の節では主にモデル2の結果を取り上げる。

■表5: モデルの適合度指標の比較

| モデル  | $\chi^2$ (df) | $p$   | CFI   | TLI   | RMSEA | SRMR | AIC     | BIC     |
|------|---------------|-------|-------|-------|-------|------|---------|---------|
| モデル1 | 6.87 (8)      | .551  | 1.000 | 1.013 | .000  | .028 | 2045.03 | 2124.15 |
| モデル2 | 7.47 (10)     | .680  | 1.000 | 1.024 | .000  | .031 | 2041.63 | 2114.42 |
| モデル3 | 50.44 (12)    | <.001 | .830  | .702  | .135  | .128 | 2080.60 | 2147.06 |

### 3.3 文法知識の種類と英語運用能力との関係

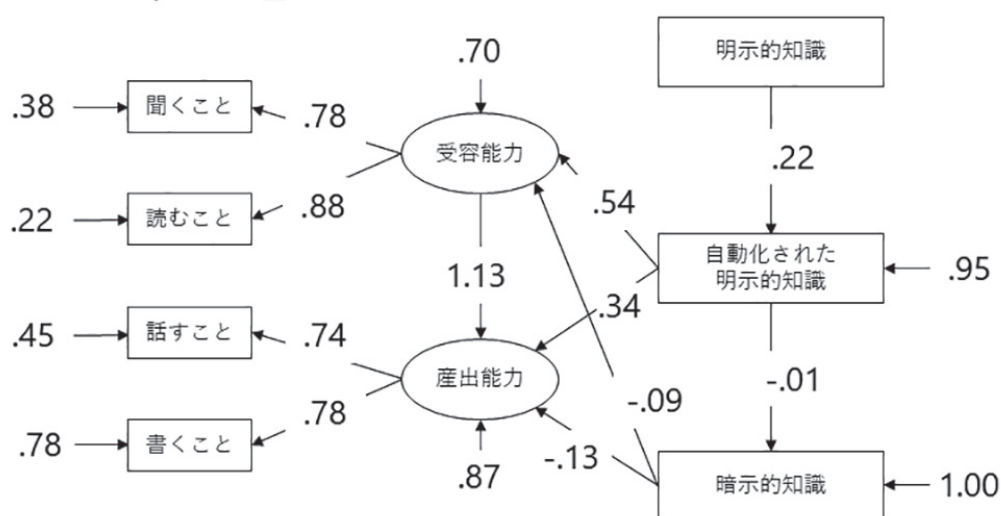
モデル2における文法知識と受容・産出能力との関係を表6、および図3に示す。このモデルは、明示的知識から自動化された明示的知識を媒介して受容・産出能力のレベルが決まるもので、明示的知識が自動化された明示的知識を有意に予測( $\beta = .22, p = .016$ )し、さらに自動化された明示的知識が受容能力( $\beta = .54, p < .001$ )、産出能力( $\beta = .34, p = .003$ )を有意に予測した。一方で、暗示的知識の寄与はモデル1と同様に、モデル2でも受容能力( $\beta = -.09, p = .457$ )、産出能力( $\beta = -.13, p = .239$ )ともに統計的に有意な影響は認められなかった。以上の結果から、自動化された明示的知識を媒介とした明示的知識の間接効果の構造が明確に現れ、技能への影響の大部分が自動化された明示的知識経由であることが明らかとなった。

■表6: モデル2における文法知識と受容・産出能力との関係

| 予測因子        | 基準変数        | 標準化係数( $\beta$ ) | $p$ 値( $p$ ) |
|-------------|-------------|------------------|--------------|
| 明示的知識       | 自動化された明示的知識 | .22              | .016         |
| 自動化された明示的知識 | 受容能力        | .54              | <.001        |
| 自動化された明示的知識 | 産出能力        | .34              | .003         |
| 暗示的知識       | 受容能力        | -.09             | .457         |
| 暗示的知識       | 産出能力        | -.13             | .239         |

注. 表中の $p$ 値は各パスの統計的有意性を示す。また、太字はその値が有意であることを示す。

## モデル 2



■ 図3: 明示的知識, 自動化された明示的知識, および暗示的知識と英語4技能との関係

注: 受容能力と産出能力はそれぞれ2指標(聞くこと・読むこと)と(話すこと・書くこと)によって測定されている。すべての数値は標準化パス係数を示す。

## 4 考察

本研究では、中学生を対象に、文法知識を明示的知識、自動化された明示的知識、暗示的知識の3種類に分類し、それぞれが英語4技能(聞くこと・読むこと、話すこと・書くこと)にどのように関与するかを検討した。その結果、3種類の知識は技能への関与の仕方や程度が異なり、自動化された明示的知識が最も安定して強い影響を示した一方で、明示的知識と暗示的知識の効果は限定的であった。

以下では、まず、構造方程式モデリングの結果から、3種類の知識が示した効果の相対的重要度を整理し、第二言語習得の理論的枠組み (Skill Acquisition Theory) や先行研究との関連からその背景を検討する。次に、それぞれの知識タイプの特徴や形成条件を、学習者の年齢や学習環境の観点から考察する。最後に、これらの知見が中学校英語教育における文法指導や練習活動の設計に与える示唆について論じる。

### 4.1 構造方程式モデリングの結果

モデル2に基づく分析の結果、自動化された明示的知識は、受容能力(聞くこと・読むこと:  $\beta = .54$ ,  $p < .001$ )と産出能力(書くこと・話すこと:  $\beta = .34$ ,  $p = .003$ )のいずれにも安定して有意な正の影響を与えていた。さらに、明示的知識は自動化された明示的知識に対して有意な正の影響を及ぼしており ( $\beta = .22$ ,  $p = .016$ )、「明示的知識 → 自動化された明示的知識 → 技能」という媒介関係が成り立つことが示された。

このことは、文法事項を「理解して知っている」だけでなく、それを実際の使用場面で素早く、しかも正確に使いこなせる状態にまで高めることが、4技能すべての向上に不可欠であることを意味する。言い換えると、知識を運用できる形に変える「自動化」が重要であるということである。この結果は、Skill

Acquisition Theory が述べる「知識は反復練習によって手続き化され、最終的に自動化される」というプロセス (DeKeyser, 2007) や, Suzuki et al. (2025) が示す「使える知識」の定義とも一致している。

一方で、暗示的知識は受容能力 ( $\beta = -0.09, p = .457$ ), 産出能力 ( $\beta = -0.13, p = .239$ ) のいずれに対しても統計的に有意な影響を示さなかった。中学生という学習初期の段階では、意味中心の豊富なインプットや自然な英語使用の経験が不足していることが多く、こうした環境では暗示的知識が十分に形成されにくい可能性がある (Ellis, 2005)。

以上の結果から、今回の参加者のような学習初期段階の学習者においては、暗示的知識よりも自動化された明示的知識を育成することが、4技能の発達に向けた、より現実的かつ効果的なアプローチであるといえる。

さらに、今回の結果は、先行研究で示されてきた文法知識の種類ごとの特徴と整合する一方で、いくつかの重要な拡張的示唆も与えている。たとえば、Zhang (2012) や Elder and Ellis (2009) では、文法知識の技能への貢献が知識の種類や対象構造によって異なることが指摘されていたが、いずれの研究も「明示的知識」「暗示的知識」の二分法に留まっていた。

まず、Zhang (2012) では「暗示的知識が読解力に対して、明示的知識よりも強く関連していた」という結果であったが、本研究の結果はこれとは異なり、受容能力に対して、暗示的知識は有意な影響を示さず、自動化された明示的知識が有意な影響を与えていたというものであった。この違いは、主に①参加者特性と②知識の扱い方の違いに起因すると考えられる。Zhang の対象は大学生の上級者であり、すでに明示的知識を習得・定着した状態で、暗示的に運用できる段階にあったと推察される。このため、読解には暗示的知識がより効果的に働いた可能性がある。一方、本研究の対象は中学生であり、暗示的知識が十分に形成されていない段階であるため、まずは明示的知識の自動化が読解力を支える役割を果たしていると考えられる。そして、Zhang は明示的／暗示的の二分法に基づいていたのに対し、本研究では自動化された明示的知識を独立した知識として捉えた。その結果として、従来は暗示的知識によるものとみなされていた処理の一部が、実際には自動化された明示的知識によるものである可能性がある。

本研究の結果は、Elder and Ellis (2009) に見られる「知識の種類と構造・受容・産出能力の対応関係」とは異なる視点を提供するものである。彼らの研究では、比較表現のような文法項目では暗示的知識が、関係詞のような文法項目では明示的知識が、それぞれ IELTS スコアとの関連を示していた。つまり、文法項目や技能の種類によって、明示的知識・暗示的知識の有効性が異なるとされていた。これに対して、本研究は、明示的知識そのものよりも、自動化された明示的知識が、聞くこと、読むこと、話すこと、書くことといった4技能すべてに安定して寄与していたことを示した。これは、文法項目や技能の種類によって知識タイプの有効性が分かれるということではなく、文法知識の「処理の自動性」が技能全体の基盤となるということを示唆する。

このように、先行研究が文法構造や産出・受容能力ごとに知識タイプの有効性が異なるとする立場を取っていたのに対し、本研究は、知識の自動化が、受容能力・産出能力の両方に関与することを示しており、文法知識と4技能との関係を再考するうえで重要な示唆を与えるものである。

また、Suzuki and Sunada (2019) が明らかにした「練習の形式や順序が関係詞の習得にどのように影響するか」という文法知識の獲得に関する知見に対して、本研究の結果は、より広い視点からの示唆を与えるものである。彼らの研究が特定の構文に対する理解や産出の正確さ・速度に注目していたのに対し、本研究では、関係詞の明示的知識が自動化されている状態が英語の4技能全体の言語運用に貢献することが示された。これは、文法知識の自動化が実際の言語使用における処理の効率や安定性に関わっていることを意味している。

さらに、本研究の対象が中学生であった点も注目に値する。発達段階にある学習者においても、明示的に学んだ知識を素早く正確に使えるようになることで、聞く・読む・話す・書くといった技能の向上に直接つながる可能性が示された。このことは、学校現場での文法指導において、知識の定着だけでなく、運用の自動化を視野に入れた練習設計の重要性を示すものである。

## 4.2 Skill Acquisition Theory との関係

本研究の構造方程式モデリングの結果、自動化された明示的知識は、聞く・読むといった受容技能だけでなく、話す・書くといった産出技能にも安定して有意な影響を与えていることが明らかになった。この結果は、Skill Acquisition Theoryにおける「知識の手続き化と自動化」という理論的枠組みと一致している。

Skill Acquisition Theoryによれば、学習の初期段階で習得された文法知識は、まず「明示的に覚えて意識的に使う知識」として始まる。その後、繰り返し練習を重ねることで「手続き的な知識」へと変化し、最終的には「意識しなくても素早く使える知識(自動化)」に到達するとされる(DeKeyser, 2007)。今回の分析では、制限時間付き文法性判断テストにおける「文法的に正しい英文の正答率」を自動化された明示的知識の指標としたが、この指標が4技能すべての成績を強く予測していた。これは、「自動化された知識が実際の言語運用を支える」という理論的主張を裏付ける結果である。

一方、明示的知識は、自動化された明示的知識の形成には有意な影響を与えていたが、それ自体が4技能のパフォーマンスに直接的な貢献を示すことはなかった。これは、明示的知識が単なるルール理解にとどまり、それが手続き化・自動化を経て初めて運用可能な知識となるという理論的枠組み(DeKeyser, 2007)と整合的である。したがって、明示的知識の習得後には、実際の使用場面で素早く活用できるようにするための練習が必要であることが示唆される(Fang, Elgort, & Chen, 2024)。

また、自動化された明示的知識と暗示的知識の関係については、Suzuki and DeKeyser (2017)の大学生を対象とした研究で、interfaceモデルにおいて 自動化された明示的知識 → 暗示的知識 のパスが有意であり( $\beta = 0.35, p < .05$ )、明示的に学習された知識が練習を通じて暗示的处理に近づく可能性が示唆された。ただし、モデル全体としてはnon-interfaceモデルとの適合度に有意な差は見られず、この移行が一貫して確認されたわけではないため、慎重な解釈が求められる。

一方、本研究の中学生データでは、この自動化された明示的知識から暗示的知識へのパスは統計的に有意ではなく( $\beta = -0.01, p = .84$ )、自動化された明示的知識から暗示的知識への移行は確認されなかった。この違いは、学習者の発達段階や学習環境の相違に起因する可能性があり、特に中学生という学習初期段階では、暗示的知識の形成に必要な豊富な意味中心のインプットや自然な使用経験が不足していると考えられる。そのため、本研究では、自動化された明示的知識が暗示的知識の前段階として機能するというよりも、直接的に4技能のパフォーマンスを支える中核的な知識として重要な役割を果たしていたことが示された。

また、関係代名詞の指導が、明示的知識を繰り返し操作し、その処理を自動化することを目的としていた点も影響した可能性がある。(2. 2) マテリアルで示した通り、実際に行われた練習課題(穴埋め・並べ替え・和訳)は、いずれも習得したルールを思い出させ、それを活性化させるように設計されていた。そのため、暗示的知識の形成は限定的となり、技能との関連が統計的に見られなかったと考えられる。

今回、4技能に対して暗示的知識の貢献が見られなかったことは、言語習得の全過程をスキル習得理論だけで説明することの限界を示す。和泉(2016)は、文法理解や練習活動に過度な期待を寄せるべきではなく、学習者が必ずしも「明示的知識 → 手続き化・自動化 → 言語習得」という直線的な経路をたどるとは限らないこと、さらに無意識的に獲得される知識(暗示的知識)やインプット主導の学習も重要であることを指摘している。

一方で、本研究の結果では、自動化された明示的知識が4技能に一貫して有意に貢献していた。この点は、スキル習得理論が想定する枠組み(明示的知識 → 手続き化 → 自動化 → 運用可能な知識)と整合している。したがって、和泉(2016)の指摘にあるように文法指導や練習に依存しすぎることは避けつつも、本研究で用いた指導方法と練習は4技能のパフォーマンス向上に一定の有効性を持つと考えられる。さらに、暗示的知識が有意な影響を示さなかったことから、その発達を支えるためには、意味のあるインプットや文脈を伴う活動を段階的に取り入れていく必要がある。

### 4.3 限界点と今後の展望

限界点としては、第一に、自動化された明示的知識の形成と各技能への貢献は一部明らかになったものの、明示的知識から自動化された明示的知識への変換過程は統計的に確認されなかった。この点は実践的・発達段階的にも重要であり、その背景には中学生段階における自動化の未成熟や、練習の質的制約が関与している可能性がある。また、Schurz (2023) が指摘するように、3種の文法知識の測定結果は課題の文脈や条件設定によって変化する可能性がある。本研究の対象は日本の私立中学校に在籍する13歳から14歳の学習者であり、英語への接触環境が授業時間にほぼ限定されていたため、結果の一般化には慎重な検討が求められる。

第二に、本研究で測定した3種の文法知識(明示的知識・自動化された明示的知識・暗示的知識)は、いずれも関係代名詞に関する文法性判断および意味性判断という特定形式の課題に基づいて指標化された。それぞれの課題が想定通りの認知処理を反映しているか、すなわち指標としての妥当性を検証する追加研究が必要である。

第三に、暗示的知識が4技能に有意な影響を示さなかったことについては、上述のような英語接触環境の制約が一因と考えられる。それにもかかわらず、このような環境下でも自動化された明示的知識が技能に有意な貢献を示したことは、教室内指導に依存する日本型英語教育に大きな示唆を与える。ただし、この知識構造や影響パターンの適用可能性をさらに確かめていく必要がある。

今後は、より多様な指標を組み合わせた測定モデルや、測定対象とする文法項目の拡充により、文法知識の妥当性を高める必要がある。また、教育実践との接続という観点からは、タスクベースの活動や即時応答型の練習が、自動化の促進および4技能のパフォーマンス向上にどのように寄与するかを検証する実証研究が求められる。特に、知識の自動化を意識した言語活動の設計と、その効果測定に関する研究は、教育現場における文法指導の最適化に大きく貢献することが期待される。さらに、本研究では対象を中学2年生に限定したが、学習者の学年や発達段階を変えて同様の分析を行うことで、知識タイプと技能の関係が年齢や学習経験の差によってどのように異なるのかを明らかにすることも、今後の重要な課題である。

### 4.4 教育的実践に対する示唆

本研究の結果、自動化された明示的知識が英語4技能すべてに有意に貢献していたことから、文法事項の理解(明示的知識)にとどまらず、それを迅速かつ正確に運用できる状態へ高める練習設計の重要性が示された。文法知識測定前の指導では、「関係代名詞の穴埋め・並べ替え・和訳課題」といった構造化された形式練習を取り入れた。これらの課題は、一斉授業の枠組みでも実施が容易であり、学習者間の取り組みの質の差に左右されにくく、学習到達度の一定水準を確保しやすいという利点を持つ。

DeKeyser (2007) は、言語知識が意識的な明示的処理から練習を通じて手続き化・自動化されるプロセスを強調し、とりわけ正確性と速度を求める繰り返し練習が自動化の促進に寄与すると述べている。本研究で採用した練習課題は、文法知識の迅速な適用を促し、意識的処理から無意識的処理への移行を支える役割を果たしていたと考えられる。

以上から、中学生のような学習初期段階の学習者においては、文法構造を用いて意味を正確に理解・表現させる課題が、明示的知識の自動化を促進し得る。暗示的知識が十分に形成されていない段階では、自動化された明示的知識の育成こそが、4技能の発達に向けた現実的かつ効果的なアプローチとなるのではないかと。

本研究は、中学生段階の英語学習において「自動化された明示的知識」の育成が重要であるという仮説に実証的根拠を与えるとともに、高校を含む教育現場に対しても、理論的に妥当かつ具体的な指針を提供するものである。

## 5 結論

本研究では、文法知識を「明示的知識」「自動化された明示的知識」「暗示的知識」の三つに分け、それぞれが英語の受容技能(聞く・読む)と産出技能(話す・書く)にどう関わるかを、中学生を対象に構造方程式モデリングで検討した。

分析の結果、自動化された明示的知識が両方の技能に最も強く、安定した影響を及ぼすことが明らかになった。特に聞く・読むといった理解面では効果が大きく、学習者が文法知識を素早く正確に使えるようになることが理解力の向上につながることが示された。明示的知識そのものは技能に直接の効果を示さなかったが、自動化された明示的知識を形成することで間接的に技能を支えていた。一方、暗示的知識は今回の中学生では有意な効果を示さず、限定的であった。

これらの結果は、スキル習得理論が示す「知識が練習を通じて自動化され、実際の言語使用に活かされていく」という段階モデルを部分的に裏付けるものである。特に、自動化された明示的知識が実際の言語活動で中核的な役割を担っていることが実証的に確認された。また、自動化の進み方が一様ではないことや、暗示的知識との結び付きが見られなかったことから、練習方法や課題の性質、学習者の発達段階が大きな影響を及ぼすと考えられる。

教育的には、単に文法を理解させるだけでなく、それを素早く使えるようにする練習が重要である。特に、明示的知識を自動化させる工夫は、4技能の向上を支える効果的な手立てとなるだろう。一方で、暗示的知識の発達をどう促すか、またそれをどう測定するかは、今後の課題として残されている。

本研究は、中学生を対象に3種類の文法知識と技能との関係を検証した初めての試みであり、文法指導と技能育成を結び付ける理論的な枠組みを提供するとともに、教室での具体的な指導に役立つ示唆を示した。今後は、年齢やレベル、課題の種類を広げた研究を通じて、知識と技能の関係をさらに深めていくことが期待される。

### 謝辞

本研究の遂行にあたり、格別のご指導と温かいご助言を賜りました千葉大学の西垣知佳子先生、神戸市外国語大学の濱田彰先生に、心より深く御礼申し上げます。また、本研究に貴重な機会をお与えくださいました公益財団法人 日本英語検定協会の皆様、ならびに選考委員の先生方に対し、謹んで感謝の意を表します。

## 引用文献

- Bowles, M. A. (2011). Measuring Implicit and Explicit Linguistic Knowledge. *Studies in Second Language Acquisition*, 33(2), 247-271. <https://doi.org/10.1017/s0272263110000756>
- Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Salience in Second Language Acquisition: Physical Form, Learner Attention, and Instructional Focus. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01284>
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Blackwell.
- DeKeyser, R. M. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 97-113). Routledge.
- Elder, C., & Ellis, R. (2009). Implicit and explicit knowledge of an L2 and language proficiency. In R. Ellis, S. Loewen, C. Elder, et al. (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 167-193). Multilingual Matters.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54(2), 103-138. <https://doi.org/10.1111/j.1467-9922.2004.00255.x>
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27(2), 141-172. <https://doi.org/10.1017/S0272263105050096>
- Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics*, 27(3), 431-463. <https://doi.org/10.1093/applin/aml022>
- Ellis, R., & Roever, C. (2018). The measurement of implicit and explicit knowledge. *The Language Learning Journal*, 49(2), 160-175. <https://doi.org/10.1080/09571736.2018.1504229>
- Fang, N., Elgort, I., & Chen, Z. (2024). Effects of retrieval schedules on the acquisition of explicit, automatized-explicit, and implicit knowledge of L2 collocations. *Studies in Second Language Acquisition*, 1-23. <https://doi.org/10.1017/s0272263124000184>
- Gass, S. (1982). From theory to practice. In M. Hines & W. Rutherford (Eds.), *On TESOL '81* (pp. 129-139). Teachers of English to Speakers of Other Languages.
- Godfroid, A., Loewen, S., Jung, S., Park, J.-H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge. *Studies in Second Language Acquisition*, 37(2), 269-297. <https://doi.org/10.1017/S0272263114000850>
- Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443-482). John Benjamins.
- Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research*, 2(1), 1-23. <https://doi.org/10.1177/136216889800200102>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hui, B. (2019). Processing variability in intentional and incidental word learning. *Studies in Second Language Acquisition*, 42(2), 327-357. <https://doi.org/10.1017/s0272263119000603>
- Hulstijn, J. H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, 18(3), 193-223. <https://doi.org/10.1191/0267658302sr207oa>
- Hulstijn, J. H., Van Gelderen, A., & Schoonen, R. (2009). Automatization in second language acquisition: What does the coefficient of variation tell us? *Applied Psycholinguistics*, 30(4), 555-582. <https://doi.org/10.1017/s0142716409990014>
- 和泉 伸一. (2016). 『第2言語習得と母語習得から「言葉の学び」を考える』. アルク.
- 小林 雄一郎, 濱田 彰, & 水本 篤. (2020). 『Rによる教育データ分析入門』. オーム社.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Li, S., Ellis, R., & Kim, J. (2018). The influence of pre-task grammar instruction on L2 learning: An experimental study. *Studies in English Education*, 23(4), 831-857. <https://doi.org/10.22275/see.23.4.03>
- Lim, H., & Godfroid, A. (2015). Automatization in second language sentence processing: A partial, conceptual replication of Hulstijn, Van Gelderen, and Schoonen's 2009 study. *Applied Psycholinguistics*, 36(5), 1247-1282. <https://doi.org/10.1017/S0142716414000137>
- 文部科学省. (2015). 英語教育の在り方に関する有識者会議 (第7回) 配付資料. [https://www.mext.go.jp/component/b\\_menu/shingi/giji/\\_icsFiles/afieldfile/2015/03/25/1356122\\_02.pdf](https://www.mext.go.jp/component/b_menu/shingi/giji/_icsFiles/afieldfile/2015/03/25/1356122_02.pdf) (2025年08月22日 閲覧)
- Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. Routledge.
- 日本英語検定協会. (n.d.). 英検の特長. <https://www.eiken.or.jp/eiken-iba/features.html> (2025年08月22日 閲覧)
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528. <https://doi.org/10.1111/0023-8333.00136>
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249-255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>

## 引用文献

- Schurz, A. (2023). Measuring teenage learners' automatized, explicit, and/or implicit knowledge: A question of context? *Language Learning*, 74(2), 506-541. <https://doi.org/10.1111/lang.12624>
- Spada, N., & Tomita, Y. (2010). Interactions Between Type of Instruction and Type of Language Feature: A Meta - Analysis. *Language Learning*, 60(2), 263-308. <https://doi.org/10.1111/j.1467-9922.2010.00562.x>
- Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, 38(5), 1229-1261. <https://doi.org/10.1017/S014271641700011X>
- Suzuki, Y. (Ed.). (2023). Practice and automatization in second language research: *Perspectives from skill acquisition theory and cognitive psychology* (1st ed.). Routledge. <https://doi.org/10.4324/9781003414643>
- Suzuki, Y., & DeKeyser, R. (2017). The Interface of Explicit and Implicit Knowledge in a Second Language: Insights From Individual Differences in Cognitive Aptitudes. *Language Learning*, 67(4), 747-790. <https://doi.org/10.1111/lang.12241>
- Suzuki, Y., & Sunada, M. (2019). Dynamic Interplay between Practice Type and Practice Schedule in a Second Language. *Studies in Second Language Acquisition*, 42(1), 169-197. <https://doi.org/10.1017/s0272263119000470>
- Suzuki, Y., Maie, R., & Hui, B. (2025). Research timeline: Automatization in second language learning. *Language Teaching*, 1-20. <https://doi.org/10.1017/s026144482500059x>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House.
- Tode, T. (2007). Durability problems with explicit instruction in an EFL context: the learning of the English copula be before and after the introduction of the auxiliary be. *Language Teaching Research*, 11(1), 11-30. <https://doi.org/10.1177/1362168806072398>
- Uchihara, T., Saito, K., Kurokawa, S., Takizawa, K., & Suzukida, Y. (2024). Declarative and automatized phonological vocabulary knowledge: Recognition, recall, lexicosemantic judgment, and listening-focused employability of second language words. *Language Learning*. <https://doi.org/10.1111/lang.12668>
- Zhang, D. (2012). Vocabulary and grammar knowledge in second language reading comprehension: A structural equation modeling study. *The Modern Language Journal*, 96(4), 558-575. <https://doi.org/10.1111/j.1540-4781.2012.01398.x>

## 第37回 研究助成

## A 研究部門・報告Ⅱ・英語能力テストに関する研究

# 高等学校の観点別学習状況の評価における 教室内英語テストの妥当性検証と改善

研究者:大鋸 雄介 神奈川県／神奈川県立七里ガ浜高等学校 教諭  
東京都／東京学芸大学連合大学院 在籍

《研究助言者:小泉 利恵》

## 概要

本研究では、Kane(1992)が提唱し、Chapelle et al. (2008), Chapelle and Voss(2021)が発展させた論証に基づく妥当性検証アプローチ(argument-based approach to validation)を用いて、教室内英語テストの妥当性を、七つの推論のうちの領域定義推論(Domain Definition Inference)及び波及効果推論(Consequence Implication Inference)の観点から検証することを目的とした。具体的には、筆者の勤務校において、英語科教師への質問紙調査、生徒への質問紙調査、言語テスト分野の専門家による文書レビューを実施した。その結果、教師は授業や教科書との整合性を重視してテストを設計している一方で、単元計画やCan-doリストとの接続には課題が見られること、構成概念への理解にばらつきがあることが明らかとなった。また、生徒質問紙調査により、定期試験がリーディングや語彙・文法の領域に対する学習動機づけに一定の影響を及ぼしていることが示された一方で、リスニングについては波及効果が限定的であり、配点や出題の有無が学習行動に影響している可能性が示唆された。これらの知見は、論証に基づく妥当性検証における領域定義推論及び波及効果推論の前提を支持する証拠の一部を提示しつつも、教室内テストに内在する課題を提示するものであり、高等学校における教室内英語テストの妥当性論証の構築に寄与するものである。一方で、その他の推論に関する証拠は本研究の範囲外であり、それらを今後の研究で補完することで、より包括的な妥当性検証が達成されと考えられる。

## 1 はじめに

平成29年・30年告示学習指導要領では、各教科等の目標及び内容を、「知識及び技能」「思考力、判断力、表現力等」「学びに向かう力、人間性等」の三つの柱で再整理した(文部科学省, 2017a, 2017b, 2018)。観点別学習状況の評価(以下、観点別評価)とは、目標に準拠した評価を実施するために、「知識・技能」「思考・判断・表現」「主体的に学習に取り組む態度」の三つの観点で学習状況を分析的に捉えることであり、これらを総括して評定を定めることとなった(国立教育政策研究所, 2020a, 2020b, 2021)。外国語科においては、小・中・高等学校で一貫した目標を実現するために、ヨーロッパ言語共通参照枠(CEFR)を参考に、五つの領域で外国語科の各科目の目標を設定し、その目標を実現するために行う言語活動の内容もCEFRを参照するという方針が示された(文部科学省, 2017b, 2018)。この方針に基づいて、外国語科では五つの領域それぞれにおいて設定された目標の実現の状況を三つの観点ごとに適切に評価するために、五つの領域についてそれぞれ三つの観点で評価規準を設定することになった(国立教育政策研究所, 2020a,

2020b, 2021)。さらに、約10年ごとに行われる学習指導要領の改訂に向けた中央教育審議会特別部会の議論では、現行の3観点のうち「主体的に学習に取り組む態度」について、評価の妥当性や現場負担の観点から評定に直接反映させない方向が検討されている一方、「知識・技能」及び「思考・判断・表現」は次期学習指導要領においても引き続き評価観点として用いられる見込みである(中央教育審議会, 2025)。

しかし、外国語科における観点別評価には、以前からさまざまな課題があると指摘されてきた。村越・江原(2015)は、ペーパーテスト作成時における各設問の観点分類に問題があることや、教師間で分類に関する認識にずれがあることを明らかにしている。この調査は旧学習指導要領に基づく評価を扱ったものであるが、平成29年・30年告示学習指導要領下においても課題が見られる。今井(2023)や馬場(2023)は、国立教育政策研究所の『「指導と評価の一体化」のための学習評価に関する参考資料』において、「知識・技能」と「思考・判断・表現」の評価規準作成時の留意点が類似している点を挙げ、「思考・判断・表現」の観点の評価の難しさを指摘している。

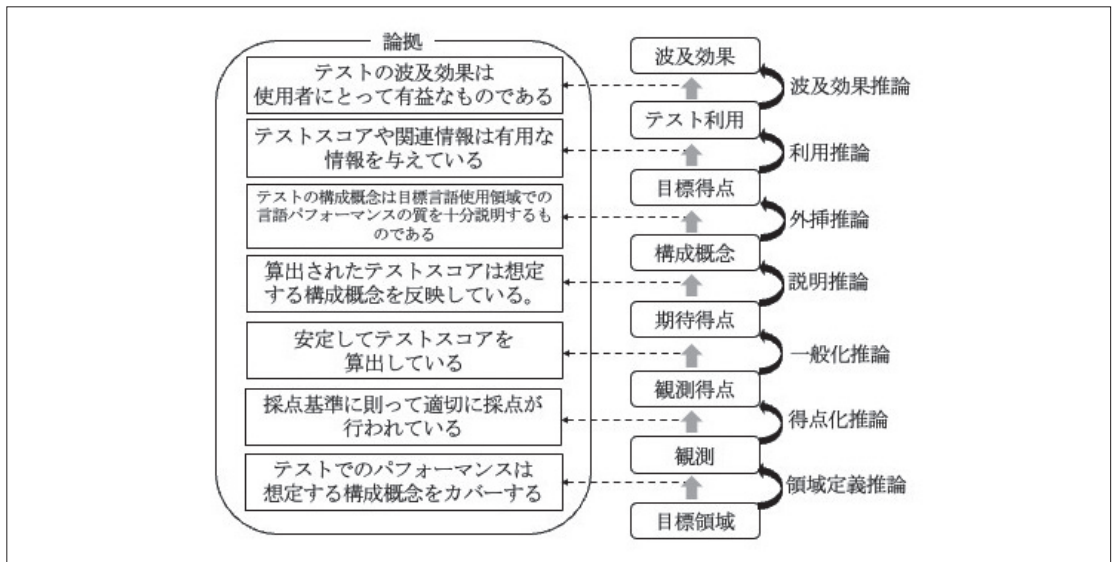
その結果、教師は何に基づいて評価をすればよいのか分からず、同様に生徒もどのように学習すればよい評価を得られるのかが分からず、その結果、生徒の学習に負の波及効果を及ぼすことが懸念される。そのような中で、ペーパーテストのみならず、幅広い方法で生徒の学習状況を上記の観点別に評価するように求められる学校現場は非常に混乱しているように感じられる。勤務校でもこの方針に則り、定期試験及びパフォーマンステストを通じた評価を行ってきたが、特に受容技能を中心とする定期試験において、どの構成概念を測定しているのかが曖昧であるという課題がある。

しかし、外国語科の観点別評価が、生徒の外国語学習の到達度を測る評価として妥当性が高いかを実証的に検証した研究は調べた限り見つからない。高等学校の評価は大学入試にも幅広く活用され、生徒の人生を左右する重要な評価であるといえる。そのため、外国語科における観点別評価の妥当性を検証することは喫緊の課題である。本研究は、言語テストにおける妥当性を検証するための「論証に基づく妥当性検証アプローチ(argument-based approach to validation)」を用いて、高等学校における教室内英語テストの妥当性を検証することを目的とする。

## 2 先行研究

妥当性は構成概念だけでなく、外部基準との関連性や波及効果、信頼性など、さまざまな要素で構成される概念である(Messick, 1996)。そのため、言語テストの妥当性を高めるには、これら複数の要素をどの程度満たしているかを多角的な手法を用いて検証する必要がある。このような妥当性の検証手法として、「論証に基づくアプローチ(argument-based approach)」が提唱された(Kane, 1992)。このアプローチでは、テストスコアに基づいて妥当性を主張するために必要な証拠を体系的に提示することで、妥当性の各要素を支持することが目指される(小泉, 2018)。特に、Kane(1992)のアプローチを修正・発展したChapelle et al.(2008)の「論証に基づく妥当性検証の枠組み(validation framework of argument-based approach)」は妥当性を論じた多くの文献で引用され、言語評価の妥当性検証にも使われている。

この枠組みにおいては、以下の六つの推論を通じてテストの妥当性を検証する:(1)領域定義推論、(2)得点化推論、(3)一般化推論、(4)説明推論、(5)外挿推論、(6)利用推論。Chapelle and Voss(2021)では上記六つの推論にテストの波及効果についての要素を含む(7)波及効果推論を加えて七つの推論とし、それらを裏付ける前提や証拠を示しながら推論の正当性を立証していく(図1)。



■図1: Chapelle and Voss (2021)の論証に基づく妥当性検証の枠組み

注:小泉(2018), 久保(2022)に基づいて著者が作成

論証に基づく妥当性検証では、まず「学習到達度の測定や指導改善への活用といった目的で実施される高等学校の定期試験が、意図した能力を適切に測定し、その使用が有益である」というような主張を立てた上で、領域定義から波及効果に至る各推論に対応する前提や反論、必要な証拠を含んだ論証を構築し、証拠の収集と論証の見直しを繰り返しながら妥当性を検証していく(小泉, 2025)。Chapelle and Voss (2021)による論証に基づく妥当性検証の枠組みでは、七つの推論を通じて言語テストの妥当性を多角的に検討する。

(1) 領域定義推論では、テストで測定すべき言語使用の文脈(目標言語使用領域: Target Language Use [TLU] domain)が適切に定められているか、またテストがそのTLUを適切に反映しているかを検証する。TLU領域とは、テストの場面とは異なる現実の場面において、受験者が実際に英語を使って何らかのやりとりやタスクを行うことが求められるような、具体的な状況や環境のことを指す(Bachman & Palmer, 2010)。例えば、高等学校の教室テストが学習指導要領に示された言語活動や学習目標に基づいて設計されているかどうかが問われる。

(2) 得点化推論では、テストによって観測された受験者のパフォーマンスが適切に採点されているかどうか論点となる。例えば、多枝選択式のテストに対しては、項目困難度や弁別力の観点から、スピーキングやライティングなどのパフォーマンステストに対しては、ルーブリックの明確性や採点の一貫性などの観点から検証する。

(3) 一般化推論では、観測得点(採点されたテストスコア)が異なる実施日や異なる採点者でも安定して得られるかを分析し、信頼性について検証する。

(4) 説明推論では、期待得点(目標領域で得られた信頼性の高いテストスコア)が、想定する構成概念をどの程度反映しているかを検証する。

(5) 外挿推論では、テスト得点がTLU領域における実際のパフォーマンスをどの程度予測・反映するかを検証する。

(6) 利用推論では、テストスコアや関連情報が意思決定のために有益に活用されているかを検証する。大規模テストにおいては、受験者を適切に分類することやカットスコアの正当化が焦点となるが、本研究の対象である教室内テストでは、教師がスコアレポートを用いて学習指導の改善や生徒へのフィードバックを行うといった教育現場での活用が主な検討対象となる。

(7)波及効果推論では、テストが学習や指導にどのような影響を及ぼしているか、例えば「定期試験が文法暗記偏重の学習を促してしまっていないか」や「スピーキングテストが生徒の発話量を増やすきっかけとなっているか」といった教育的影響を考察する。

これらの推論を統合的に検討することで、テストの設計・実施・運用に関する妥当性の全体像を論証的に構築することが可能となる。以下の表1は、上述した妥当性論証における主要な推論・前提・証拠の例をまとめたものである。

■表1: 言語テストの妥当性論証における主要な推論・前提・証拠の例 (Dai et al., 2024に基づく)

| 推論                                                       | 前提の例                                                                                                                                                                                           | 証拠の例                                                                                                                                                       |
|----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 領域定義推論<br>主張: テストタスクの選定・設計・実施が、関連する目標領域 (TLU 領域) を考慮している | <ul style="list-style-type: none"> <li>・テストタスクが目標領域を反映している</li> <li>・テストタスクが目標領域を十分に捉えている</li> </ul>                                                                                           | <ul style="list-style-type: none"> <li>・領域分析, 領域関係者へのインタビュー及び質問紙調査</li> </ul>                                                                              |
| 得点化推論<br>主張: 受験者のパフォーマンスが、意図された特性をもつ得点として適切に評価されている      | <ul style="list-style-type: none"> <li>・実施手順が適切である</li> <li>・項目の統計的特性は適切である</li> <li>・テストタスクは受験者を適切なレベルに分類できる</li> <li>・テストタスクは受験者が能力を発揮する十分な機会を提供している</li> <li>・採点には信頼性がある</li> </ul>         | <ul style="list-style-type: none"> <li>・テスト条件の検討, 受験者へのインタビュー</li> <li>・統計分析 (項目分析, Rasch 分析)</li> <li>・関係者へのインタビュー</li> <li>・応答特性を含むテスト問題のレビュー</li> </ul> |
| 一般化推論<br>主張: 観測された得点は、他の類似課題や形式、採点者にわたって再現可能なものである       | <ul style="list-style-type: none"> <li>・タスク・項目・採点者数は信頼性の高い得点を得るのに十分である</li> <li>・テスト得点の等化処理が適切である</li> <li>・構成概念とは無関係な分散が、実施条件や受験者の母語、性別といった要因によって生じない</li> <li>・採点はテストの実施ごとに一貫している</li> </ul> | <ul style="list-style-type: none"> <li>・一般化可能性理論研究</li> <li>・テスト開発文書のレビュー</li> <li>・DIF (特異項目機能) 分析, 統計的モニタリング</li> </ul>                                  |
| 説明推論<br>主張: 期待得点は、テストが測定しようとする構成概念に起因している                | <ul style="list-style-type: none"> <li>・必要とされる知識・プロセス・方略が理論と一致している</li> <li>・テストスコアの内部構造が理論と整合的である</li> <li>・タスクにおけるパフォーマンスは、関連する他のタスクにおけるパフォーマンスと関連している</li> </ul>                            | <ul style="list-style-type: none"> <li>・受験者の言語使用の発話プロトコル</li> <li>・因子分析</li> <li>・相関分析</li> </ul>                                                          |
| 外挿推論<br>主張: テストで測定された構成概念は、現実の目標領域におけるパフォーマンスを説明できる      | <ul style="list-style-type: none"> <li>・テストでのパフォーマンスが TLU のパフォーマンスと関連している</li> <li>・受験者がテストタスクで行うやりとりが TLU のパフォーマンスと似ている</li> </ul>                                                            | <ul style="list-style-type: none"> <li>・領域関係者へのインタビュー</li> <li>・テストと現実場面の発話分析</li> </ul>                                                                   |
| 利用推論<br>主張: テストスコアや関連情報は適切な意思決定のために有益な情報を与えている           | <ul style="list-style-type: none"> <li>・評価は受験者を意思決定に適切なレベルに分類できる</li> <li>・カットスコアが正当化できる</li> </ul>                                                                                            | <ul style="list-style-type: none"> <li>・統計分析</li> <li>・基準設定方法のレビュー</li> <li>・利用者へのインタビュー</li> </ul>                                                        |
| 波及効果推論<br>主張: テストの波及効果は使用者にとって有益なものである                   | <ul style="list-style-type: none"> <li>・合格者は当該領域で十分な能力を有する</li> <li>・テスト形式が当該領域の言語使用に合致している</li> <li>・テスト得点やフィードバックが学習や指導に良い影響を与える</li> </ul>                                                  | <ul style="list-style-type: none"> <li>・受験者, 雇用者, 教師, 利害関係者へのインタビュー</li> </ul>                                                                             |

Dursun and Li (2021) は、2000年から2018年にかけて発表された言語テスト分野の論証に基づく妥当性検証研究を対象としたレビュー研究であり、その使用傾向や地域的分布を明らかにしている。彼らは論証に基づく妥当性検証の枠組みが近年増加傾向にあり、これが主流な枠組みであること、また、対象となった言語テストは大規模テストが最も多く、教室内テストを対象とした研究が最も少ないことを指摘している。そして、同枠組みにおいて証拠が示された推論のうち、領域定義推論が最も少なく、次いで利用推論及び波及効果推論が少ないことも明らかとなった。また、論証に基づく妥当性検証の枠組みを用いた研究の大半は北米(特に米国)で実施されており、70の対象研究のうち56(80%)が北米の研究者によって発表されている。一方、アジア地域の研究者によるものはわずか6件であり、その中に日本の事例も含まれているものの、全体に占める割合はごく少数である。このことは、論証に基づく妥当性検証が国際的に広がりを見せているものの、日本における実証的な適用例は依然として限られていることを示しており、今後の国内研究においてこの枠組みを積極的に活用していく必要性を示唆している。

日本の言語テストを対象に、論証に基づく妥当性検証の枠組みが用いられた数少ない例として以下の二つの研究が挙げられる。Koizumi et al. (2011) は、日本人英語学習者を対象とした文法診断テスト(EDiT Grammar)に対して、論証に基づく枠組みを用いて妥当性検証を行い、Chapelle et al. (2008) の六つの推論に沿って解釈的主張を構築し、ラッシュ分析や発話プロトコル分析を通じて、得点化推論及び説明推論を支持する証拠を示している。Kumazawa et al. (2016) は日本の大学生向けに開発された英語ブレースメントテストであるVELC Test®を対象に、Kane (2006) に基づいて妥当性検証を行った。Kaneの提唱する論証に基づく妥当性検証では、採点(scoring)、一般化可能性(generalizability)、外挿(extrapolation)、決定(decision)という四つの推論に焦点が当てられ、VELC Test®に対するそれぞれの推論に対応する証拠が統計的に収集された。例えば、ラッシュモデルに基づく項目分析や一般化可能性理論(G理論)を用いた信頼性の検証、受験者のCan-doに基づくスコアマッピングなどを通じて、テストスコアが構成概念を妥当に反映していることが示された。また、スコアレポートが受験者や教師にとって学習支援のツールとして有用であることも検討されており、Chapelle and Voss (2021) における利用推論や波及効果推論につながる視点も含まれている。このように論証に基づく妥当性検証の枠組みを用いた研究において日本の言語テストを対象とした研究は少なく、特に教室内英語テストの妥当性を検証した研究は見当たらない。よって本研究では、日本の教室内英語テストの妥当性を論証に基づく枠組みを用いて検証し、その中でも領域定義推論及び波及効果推論に係る前提についての証拠を提示し、より良い教室内英語テスト開発のために寄与することを目的とする。

領域定義推論は、上述したようにテストで観察されるパフォーマンスが、想定する構成概念を十分にカバーしていることを前提とする推論である。その検証は主に二つの視点から行われる。第一に、評価が行われる文脈におけるTLU領域で測定すべき能力が明確に定義されているか、第二に、設定されたテストタスクが、そのTLU領域における能力をどの程度反映しているかである。例えば、高等学校の教室内英語テストにおいては、学習指導要領に示された言語活動や学習目標に基づき、実際の言語使用状況を反映したテストタスクが設計されているかどうかを検討対象となる。領域定義推論で用いられる研究手法として主にカリキュラム分析、授業観察、ニーズ分析、コーパス開発と分析が挙げられる。例えば、Park (2021) は、軍用航空管制官を対象とした航空英語評価の開発において、文書分析・専門家へのインタビュー・ニーズ分析を通じてTLU領域を明確にし、そのTLUを反映した評価タスクを仮想環境で設計・実施することで、領域定義推論に基づく妥当性の検証を行っている。

波及効果推論は、テストが指導や学習に与える望ましい、あるいは望ましくない影響を明らかにする推論である。波及効果推論で用いられる研究手法としては、テスト使用者へのインタビュー、分類精度と一貫性の分析、ラッシュ分離指数、縦断的データ分析が挙げられる(Dursun & Li, 2021)。Allen et al. (2024) は論証に基づく妥当性検証の枠組みは用いていないものの、生徒質問紙調査・授業観察・インタビュー等の混合的手法を通じて日本の高等学校における英語4技能外部試験の波及効果を調査し、明確な

指導・学習への影響は確認されなかったものの、当該試験がカリキュラムや授業と整合していること、そして大学入試による負の波及効果が依然として存在することを明らかにした。

本研究の妥当性論証においては、領域定義推論の前提を「高等学校における教室内英語テストにおいて、TLU領域は明確に定義されており、テストのタスクがTLU領域の重要なスキルを適切にカバーしている」とし、波及効果推論の前提を「高等学校における教室内英語テストは生徒に有益な波及効果を与えている」と設定した。これらの前提を検証するために、以下のリサーチクエスチョンを設定した。

- RQ1

高等学校における教室内英語テストにおいて、目標言語使用領域は明確に定義されているか。
- RQ2

高等学校における教室内英語テストのタスクが、目標言語使用領域の重要なスキルを適切にカバーしているか。
- RQ3

高等学校における教室内英語テストは、生徒にどのような波及効果を与えているか。

3 方法

3.1 対象となる教室内英語テストの文脈

本研究の対象校は、神奈川県内に所在する公立高等学校であり、全日制・共学の普通科高校である。生徒数は約1,100名で、英語科には11名の専任教師が勤務している。定期試験は年間5回実施されており、パフォーマンステストは学年によって異なるものの年間1～2回実施されている。定期試験では主にリーディング及びリスニングの受容技能を、パフォーマンステストではスピーキング及びライティングの産出技能を測定している。定期試験の作成は、原則として学年担当教師間で分担されており、事前に出題範囲や試験形式に関する調整が行われる。また、パフォーマンステストについても、共通の課題とループリックを用いながら、各教師が担当クラスでの実施・評価を行っている。いずれのテストも、実施の1～2週間前を目安に、生徒に対して出題形式や評価基準等の概要を事前に周知することが慣例となっている。本研究では、高等学校における教室内英語テストの妥当性検証の一部として、領域定義推論及び波及効果推論における前提を支持する証拠を収集することを目的とし、教師質問紙調査、専門家による文書レビュー、生徒質問紙調査の三つの調査を行った。各調査の概要は表2の通りである。

表2: 本研究で実施した調査の概要

| 調査名          | 対象              | 調査内容                                                            | 実施時期    | 備考                               |
|--------------|-----------------|-----------------------------------------------------------------|---------|----------------------------------|
| 教師質問紙調査      | 筆者の勤務校の英語科教師10名 | 教室内英語テスト全般についての認識・実践を調査                                         | 2025年3月 | 特定のテストに限定せず、全般的な傾向を把握            |
| 専門家による文書レビュー | 大学教授(専門家)       | ・3年生対象の定期試験(1回)<br>・3年生対象のライティングテスト(1回)<br>・2年生対象のスピーキングテスト(1回) | 2025年3月 | 各テストに関する関連文書(ループリック、単元計画等)も併せて送付 |
| 生徒質問紙調査      | 高校1～3年生         | 各学年の1学期中間試験(定期試験)についての受け止めを調査                                   | 2025年5月 | 英語コミュニケーションの試験実施後に調査             |

## 3.2 領域定義推論に関する調査(RQ1及びRQ2)

### 3.2.1 教師質問紙調査

本調査には、筆者の勤務校である神奈川県内の公立高等学校に勤務する英語科教師10名が参加した(筆者は除く)。調査対象者の英語指導歴は、0～5年が3名、6～10年が4名、11～20年が2名、21年以上が1名であり、若手からベテランまで幅広い経験年数の教師が含まれていた。質問紙調査は2025年3月に Google Forms を用いてオンラインで実施され、事前に調査の目的、回答内容の匿名性、研究へは任意参加である旨を説明し、文書による同意を得た(資料1参照)。質問項目は Jun(2014)を参考に筆者が作成した(資料2参照)。質問項目は英語指導歴と担当学年を問う項目から始まり、Q9を除く Q3～Q19 及び Q22, 23は6段階リッカート尺度(6:「とてもよく当てはまる」, 1:「全く当てはまらない」)を用いた項目であり、中間の5, 4, 3, 2についてはラベルを付与せず数値のみを提示した。Q20 及び Q21は、定期試験やパフォーマンステストを作成する際に参照する文書について尋ねる複数回答式の項目であった。Q24～Q27は定期試験やパフォーマンステストを作成・実施・採点する際の困難点や工夫している点について自由記述回答を求める項目である。Q9は「ご自身の勤務校では Can-do リストを設定していますか?」という質問に対して「設定している」「設定していない」「分からない」の三つから選択する項目であった。自由記述を含む設問においては、回答者個人が特定されないように十分な配慮を行った。なお、本調査は特定の定期試験やパフォーマンステストに対して行われたものではなく、昨年度各教師が実施した定期試験やパフォーマンステストを作成・実施・採点する際の意識を全般的に調査したものであることや、本研究における結果の解釈及び一般化に際しては、同一校内の教師のみを対象とした調査であることに留意する必要がある。

### 3.2.2 専門家によるフィードバック

2025年3月、言語テストを専門とする大学教授に対し、勤務校で実施された1回分の定期試験、スピーキングを測るパフォーマンステスト及びライティングを測るパフォーマンステストをそれぞれ1回分、並びに関連文書(単元の指導と評価の計画、ルーブリックや Can-do リスト、教材のコピー等)を添付し、評価の妥当性に関するフィードバックをメールにて依頼した(送付した文書の一部は資料3, 4参照)。送付した文書の一覧は以下の通りである(表3)。フィードバックを行う際には、以下の三つの観点でフィードバックをいただくよう依頼した:①テスト問題が、TLU領域を適切に反映しているかについて、②評価基準(ルーブリック)の適切さと明確さについて、③試験全体の妥当性や改善すべき点について。翌月の2025年4月に、当該教授より文書によるフィードバックを受領した。フィードバックの対象となった定期試験の概要は表4の通りである。

■表3: 専門家に送付した文書一覧

| 分類                | 文書名及び内容の詳細                                                       | テスト実施時期 | 対象学年  | 作成者                      |
|-------------------|------------------------------------------------------------------|---------|-------|--------------------------|
| リスニング・リーディングテスト関連 | 英語コミュニケーションⅢ<br>1学期期末試験                                          | 2024年7月 | 高校3年生 | 筆者                       |
| スピーキングテスト関連       | 英語コミュニケーションⅡ Lesson 4<br>ゴールタスク(スピーキングタスク)                       | 2024年7月 | 高校2年生 | 勤務校英語科教師<br>(経験年数6年～10年) |
| ライティングテスト関連       | 論理・表現Ⅲ Lesson 9<br>ゴールタスク(ライティングタスク)                             | 2024年9月 | 高校3年生 | 筆者                       |
| その他<br>参考資料       | ・勤務校の Can-do リスト<br>・各科目の「単元の指導と評価の計画」<br>・該当単元の教科書及びワークシートの PDF | —       | —     | —(該当科目全体の資料として添付)        |

■表4: 対象となった定期試験の概要

| 大問  | 小問 | 問題概要                               | 形式    | 項目数 | 配点 | 評価の観点    |
|-----|----|------------------------------------|-------|-----|----|----------|
| 大問1 | A  | 副教材からのリスニング問題(ディクテーション)            | 記述式   | 8   | 8  | 知識・技能    |
|     | B  | 副教材からのリスニング問題(内容理解)                | 多枝選択式 | 3   | 6  | 思考・判断・表現 |
|     | C  | 副教材からのリスニング問題(内容理解)                | 多枝選択式 | 3   | 6  | 思考・判断・表現 |
| 大問2 |    | 副教材からの文法・語法問題                      | 多枝選択式 | 20  | 20 | 知識・技能    |
| 大問3 |    | 副教材からの語彙問題                         | 多枝選択式 | 20  | 20 | 知識・技能    |
| 大問4 | A  | 教科書からのリーディング問題<br>(教科書をパラフリーズした問題) | 多枝選択式 | 9   | 10 | 思考・判断・表現 |
|     | B  | 教科書からのリーディング問題<br>(教科書をパラフリーズした問題) | 多枝選択式 | 5   | 10 | 思考・判断・表現 |
| 大問5 | A  | 大学入試からのリーディング問題                    | 多枝選択式 | 5   | 10 | 思考・判断・表現 |
|     | B  | 大学入試からのリーディング問題                    | 多枝選択式 | 5   | 10 | 思考・判断・表現 |

今回フィードバックの対象としたスピーキングテストは、高等学校2年生の英語コミュニケーションⅡの授業において実施された。当該単元は動物の生態をテーマとし、スピーキングテストのタスクは、各生徒が自ら選択した動物について調査し、まとめた内容を発表する発表形式のタスクであった。タスクは1種類であり、場面設定は、近隣の中学校にゲストスピーカーとして招かれ、理科の授業で動物に関するスピーチを行うというものであった。発表は授業内で行われ、一人ずつポスターを用いてスピーチを行い、その場で担当教師が事前に提示されたループブリックに基づいて採点した。教科書の単元は四つのパートから構成されており、動物に関する説明文の構成、関連語彙・文法はこのタスクに関連しており、単元全体が発表タスクの準備プロセスとして構成されていた。図2に対象となったスピーキングテストの指示文を示す。

### The Goal Task

「Animal 100 選 大図鑑！これがガ浜の生徒の総力！」

In the end of this lesson, you are going to go to KOSHIGOE junior high school as a guest speaker. You will introduce an animal in science class. You are going to choose one animal and become a master of it.

■図2: 対象となったスピーキングテストの指示文

今回フィードバックの対象としたライティングテストは、高等学校3年生の論理・表現Ⅲの授業において実施されたものである。タスクは一つで、ソーシャルメディアの利点と欠点に関する英文を読んだ上で、“Do the benefits of social media outweigh the drawbacks? Agree or disagree?”というテーマについて、80~100語程度のエッセイを作成するものであった。生徒は授業時間内に20分間でライティングを行い、終了後に提出した。採点は事前に生徒へ提示されたループブリックに基づいて授業時間外に教師が行った。授業内容との関連として、当該単元では意見文の構成や主張・理由の展開方法について事前に指導を行っており、本タスクはそれらの学習内容の到達度を評価する目的として位置づけられていた。図3は対象となったライティングテストの指示文である。

**STEP 3** Read **Input A and B** and write an essay of approximately 80-100 words about **“Do the benefits of social media outweigh the drawbacks? Agree or disagree?”** in which you clearly state your position on the issue.

■図3: 対象となったライティングテストの指示文

### 3.3 波及効果推論に関する調査(RQ3)

波及効果推論の前提に関する証拠収集の一環として、2025年5月に実施した1学期中間試験の直後に、生徒を対象とした質問紙調査を実施した。対象は筆者の勤務校に在籍する高等学校1～3年生全体( $N = 1,103$ )であり、調査はGoogle Formsを用いてオンラインで実施され、事前に調査の目的、回答内容の匿名性、研究への任意参加である旨を説明し、文書による同意を得た(資料5参照)。質問項目は、定期試験が生徒の学習行動や学習意欲に与えた影響を把握することを目的とし、Allen et al. (2024)を参考にして筆者が作成した。質問項目は、主に6段階リッカート尺度(6:「とてもよく当てはまる」、1:「全く当てはまらない」)を用い、中間の5, 4, 3, 2についてはラベルを付けず数値のみを提示した。また、選択式及び自由記述式の項目も含まれていた(資料6参照)。任意回答とした結果、924名から有効回答が得られた(回収率83.8%)。なお、調査実施にあたり、回答結果は評価や成績には一切影響を及ぼさないことを生徒に周知し、匿名性とプライバシーの保護に十分配慮した。また、本研究における結果の解釈及び一般化に際しては、同一校内の生徒のみを対象とした調査であることに留意する必要がある。

## 4 結果

### 4.1 教師質問紙調査による領域定義推論の検証

質問紙調査の質問項目、及び自由記述回答の項目を除く結果を資料2に示す。資料2より、テストで測る英語力は学習指導要領や単元の指導と評価の計画等に明示されているかという質問に対して、Q3(定期試験)及びQ4(パフォーマンステスト)の平均値はそれぞれ4.3及び4.1となり、6段階尺度(6:「とてもよく当てはまる」、1:「全く当てはまらない」)の中ではやや肯定的寄りの傾向が見られた。なお、本調査では2～5の中間値にはラベルを付与していないため、これらの数値は端点との相対的位置に基づく傾向として解釈している。実際に明示されているかどうかは、別途文書分析を通じて検証する必要があるが、概ね参加教師はTLU領域で測定すべき能力が明確に定義されていると認識していることが分かった。英語指導経験年数別に見ると、経験年数が長いほど明示されているとする傾向があった。ただし、標準偏差はいずれも約1.5と大きく、教師間で認識にばらつきがあることが示唆された。

観点別評価の評価観点の違いについての理解は、技能ごとに平均値に若干の差が見られた。Q5(読むこと: $M = 3.8$ )、Q7(書くこと: $M = 3.7$ )に対し、Q6(聞くこと: $M = 3.3$ )及びQ8(話すこと: $M = 3.1$ )はやや低い値を示した。特に「話すこと」では標準偏差が1.1と大きく、教師間で理解のばらつきが相対的に大きいことがうかがえた。

さらに、詳しく回答結果を見てみると、Q5からQ8について6段階の尺度の最高値である「6」及び最低値「1」を選択した参加者はいなかった。また、「読むこと」「書くこと」「聞くこと」は6段階中の「4」を選択した参加者が最も多かったのに対し、「話すこと」については6段階中「2」を選択した参加者が最も多かった。本調査では実際にどのように評価観点を理解しているかを調査していないため、正しく理解しているかどうかは判断できない点について留意すべきではあるが、評価の根拠となる目標や観点が曖昧であり、理解が教師によって異なる可能性が示唆された。

なお、定期試験及びパフォーマンステストの設計において、教科書との整合性については高い認識が共有されていた( $Q15 = 4.6$ ,  $Q19 = 4.7$ )。また、Q22及びQ23ではいずれも平均値が5.0であり、授業内容とテストとの整合性についても高い認識が得られた。これにより、参加者は教室内テストを授業内容や教科書を反映させるものとして設計していることが示唆された。一方で、「学習指導要領」、「Can-doリスト」や「単元の指導と評価の計画」との整合性については、Q12( $M = 3.7$ )、Q13( $M = 3.5$ )、Q14( $M = 3.6$ )、Q16( $M = 4.3$ )、Q17( $M = 3.7$ )、Q18( $M = 3.9$ )に見られるように、パフォーマンステストと学習指導要領の整合性への意識は一定程度見られるものの、教科書や授業内容との整合性と比べて相対的にやや意識が低い傾向も確認された。

さらに、Q20「定期試験の設計プロセスにおいて、どのような資料を参照していますか？」及び Q21「パフォーマンステストの設計プロセスにおいて、どのような資料を参照していますか？」の回答結果を表5に示す。いずれのテストにおいても「教科書」(定期試験:90%, パフォーマンステスト:90%)が最も多く参照されており、次いで「過去の定期試験」(同90%, 70%)や「英検などの外部試験」(同80%, 70%)といった既存の問題例を参考にする傾向が強く見られた。一方で、「学習指導要領」(同10%, 0%)や「単元の指導と評価の計画」(同30%, 30%), 「Can-do リスト」(同10%, 30%)など、制度的な評価枠組みの参照は相対的に少なかった。これらの結果は、テスト設計が教科書や過去問題、外部試験といった資料に強く依拠している一方で、学習指導要領や Can-do リストなどの制度的文書との接続が十分でない可能性を示唆している。

また、Q9「ご自身の勤務校では Can-do リストを設定していますか？」の質問に対して、「設定している」が6名、「設定していない」が2名、「分からない」が2名であった。Can-do リストについては各校で作成し、活用することが推奨されており、勤務校でも独自の Can-do リストは作成してあったが、その認知度には個人差があることが分かった。このことから、一部の教師にとっては Can-do リストが形骸化している可能性が示唆された。

■表5: Q20及びQ21への回答結果(複数回答可)

| 参考とする資料     | Q20:定期試験 | Q21:PT  |
|-------------|----------|---------|
|             | 回答件数(%)  | 回答件数(%) |
| 学習指導要領      | 1 (10%)  | 0 (0%)  |
| 単元の指導と評価の計画 | 3 (30%)  | 3 (30%) |
| Can-do リスト  | 1 (10%)  | 3 (30%) |
| 教科書         | 9 (90%)  | 9 (90%) |
| 過去の定期試験     | 9 (90%)  | 7 (70%) |
| 英検などの外部試験   | 8 (80%)  | 7 (70%) |
| 大学入試過去問     | 1 (10%)  | —       |
| その他の問題集     | 1 (10%)  | —       |
| 参考となる写真、動画  | —        | 1 (10%) |

注. PT = パフォーマンステスト

自由記述回答内容については質的に分析し、意味のまとまりごとにコード化した上で、類似の内容を基にカテゴリに整理した。まず、Q24の「定期試験を作成する際に難しいと感じる点」に対する回答について、教師が定期試験を作成する際に直面している主な課題として、観点別評価における「思考・判断・表現」の観点に対応した問題作成の難しさが指摘された。また、単元の指導と評価の計画に即した問題を作成する時間的余裕の不足も挙げられた。このことは、上記の単元の指導と評価の計画との整合性の低さを示した結果の理由を説明する回答であった。さらに、テストの平均点の調整の難しさ、複数教師による担当のため試験方針にばらつきが生じやすいこと、マーク式の導入による記述問題の制限、教科書の単なる暗記に偏らない試験作りに苦慮している実態も明らかとなった。これらは、村越・江原(2015)の教師質問師調査の結果と類似している点が多く、旧学習指導要領下における定期試験作成に関する困難点は、平成29年・30年告示学習指導要領下でも解消されていない可能性が示唆されている。

次に、Q25の「パフォーマンステストを作成・実施する際に難しいと感じる点」に対して、大きく分けて時間的制約と評価の信頼性に関する回答を得た。時間的制約に関しては、40名規模のクラスで実施する場合、授業回数を複数回費やす必要があり、他の内容を扱う時間が確保しづらいという意見が複数見られた。本調査は特定のパフォーマンステストについて調査しておらず、実際の手順は不明であるが、筆者の勤務校では、スピーキングのパフォーマンステストを行う際、しばしば生徒一人ずつを廊下に呼び、パフォーマンスを行わせ、授業者が評価・フィードバックを行うことが慣例となっている。そのため、50分で1時間分の授業とすると、一度のパフォーマンステストで2~3時間費やすことになる。また、トピックや形式の検討に時間がかかることや、記述式やスピーキング活動において準備・実施・採点にかかる負担の大き

さも指摘された。小泉(2022)は、授業中に発話を録音して提出し、後で採点するなど1回の授業内で実施が可能であるスピーキングテストの実施形態を紹介しており、妥当性、信頼性、実現可能性の三つの要素から学校環境に応じた実施形態を検討することを推奨している。したがって、時間的制約の課題を軽減するためには、こうした代替的な実施形態の導入を検討することが有効であると考えられる。評価の信頼性に関しては、共通のルーブリックを用いたとしても担当者によって評価にばらつきが生じることや、フィードバックの在り方、評価者の主観を避ける難しさなど、評価基準の共通理解や指導と評価の一体化に関わる課題が多く挙げられた。その他、生徒によってモチベーションが異なる中での到達目標の設定や、テストのモデルを事前に示すことにより回答が似たものになってしまう点も挙げられた。

Q26の「定期試験やパフォーマンステストを作成・実施する際に工夫している点」に対しての回答は、大きく授業との整合性と学習者支援・動機づけの二つに分類できる。まず、授業との整合性に関しては、多くの教師が、教科書や副教材、授業で扱った内容に即した出題を意識しており、生徒が授業で得た知識や技能を試験で活用できるように設計している。例えば、語彙や文法をそのまま出題するのではなく、言い換えや応用的な文脈で問うことで、表層的な暗記では対応できないような問題に工夫している。また、試験範囲や出題形式、評価項目を事前に明示し、学習の見通しを持たせる工夫も見られた。

一方で、学習者支援と動機づけの観点では、大学入試とは直接関係のない活動であっても、英語がコミュニケーションの手段であることを意識させ、学ぶ意味や目的を丁寧に伝えることで、生徒の内発的動機づけを促す工夫が見られた。特に話す活動に対して不安を感じやすい生徒に向けては、安心して取り組めるような説明やモデル提示、練習機会の提供を通して、生徒の不安軽減を図る工夫が共有された。また、大学入学共通テストなどの大学入試に似た形式の問題を導入することで、生徒の関心を引き、学習の意欲につなげている事例もあった。この取り組みは、廣森・小金丸(2024)が指摘するように、生徒が有している学習動機を意識的に捉え、それを授業や評価の設計に反映することでエンゲージメントを促進する有効な事例であると位置づけられる。

最後に、Q27の評価を実施する際の疑問点に関する質問に対しては、A～Cの観点別評価から5段階の評定への変換の妥当性や、主体的学習に取り組む態度の評価方法に対する疑念が見られた。また、ライティング課題における生徒のAI使用やスピーキング評価の基準に関する悩みも複数報告された。こうした声は、学習評価を取り巻く環境の変化やツールの進展に対する対応の難しさを反映しており、今後の制度設計や研修の在り方に関わる重要な示唆といえる。

## 4.2 専門家フィードバック

### 4.2.1 目標言語使用領域の適切性について

専門家は、定期試験及びパフォーマンステストがTLU領域を適切に反映しているかを検討した。はじめに、テストにおいて何をTLU領域と定義しているかによって適切かどうかが変わるという指摘があった。勤務校ではテストごとにテスト仕様書(Test Specifications)を作成しているわけではなく、TLU領域がテストごとに明示されているわけではない。テスト仕様書とは、テストの目的、測定したい能力、問題の形式や内容、項目数や配点、難易度や出題範囲などを明確に定めたテストの設計図のような文書であり、評価の信頼性や妥当性の向上に寄与する(Jin, 2021; 小泉ら, 2017)。そのため、領域定義推論の観点から、改善点としてテストの実施の際はテスト仕様書を作成し、TLU領域を明示することが考えられる。

学習指導要領に基づいて「高校生にとっての日常的、社会的な英語」のようにTLU領域を定義したと仮定すれば、筆記の定期試験、パフォーマンステストともねらいとするTLU領域と、テストで測ろうとしている言語使用の領域は概ね合致しているとのフィードバックを受けた。しかし、送付した定期試験の大問5A(大学入試過去問)については、場面や目的の設定が乏しく、実際の言語使用状況との対応が弱いとの指摘があった(図4参照)。この大問5Aにおいては、問題の指示文がRead the passage below and answer the following questions.であり、目的や場面、状況の設定が欠如している。国立教育政策研究所(2021, p. 65)の参考資料が示す「読むこと」の評価事例では、You are asked to read the following

article about the Universal Design Font (UD font) to prepare for a presentation in English class. という状況設定が記載されている。さらに、大学入学共通テスト英語(リーディング)においては、Your host family in the US is planning to buy a small aquarium. You found this helpful pamphlet. など という状況設定が提示されている(大学入試センター, 2025a, p. 4)。これらを参考に、今回フィードバックを依頼した大問5の例においては、You are asked to read the following article about Irene Sendler to prepare for a presentation in English class. などのような記述を冒頭に加えることが考えられる。目的や場面、状況の明示における解釈についてはさまざまな議論があるが(渡邊ら, 2025)、テストを実施する際に、TLU 領域を明確に提示することだけでなく、その言語使用領域との対応を明示するために、問題項目に言語使用の目的や場面、状況を明記することの重要性が示唆された。

5【応用問題】<マークシート問題> 思考・判断・表現[20]

[A] Read the passage below and answer the following questions.

1 In 1942, as Jews throughout Europe were being rounded up and transported altogether to Nazi \*concentration camps to face an unthinkable fate, one woman took courageous action and risked her own life to save thousands of others.

2 Her name was Irene Sendler. Though she herself was not Jewish, she feared for the lives of Jews around her, particularly the children. As head of the children's section in the Polish underground movement known as Żegota, she was unable to sit by and not do anything. So, she went into the Warsaw Ghetto and persuaded Jewish parents and grandparents to place their children in her care, saying that they were certain to die in the Ghetto or in the Nazi death camps unless they could be spirited away to safety.

3 \*Smuggling the children past Nazi guards through a variety of means — hiding them in body bags or under loads of goods — Ms. Sendler took them into the homes of Polish families, where they were adopted and raised with false identities. Ms. Sendler made lists of these children and placed the lists in a jar that she buried in a garden, hoping she could someday dig up the jar, locate the children and inform them of their past.

4 From 1942 to 1943, Ms. Sendler managed to smuggle 400 children out of the Warsaw Ghetto before she was captured by the Nazis and severely injuring for her actions. Even under extreme torture, she refused to reveal where the lists of the smuggled Jewish children were hidden. Eventually, a member of the Polish underground bribed a guard to release her, and she entered into hiding. Even then, she continued to work with Żegota to rescue another 2,100 children.

5 It's a remarkable story and considering all the remarkable stories from the Holocaust that have surfaced over the past several years, it's hard to believe that this one went largely unnoticed. And it might have stayed that way, were it not for four high school students from Uniontown, Kansas.

\*concentration camps 強制収容所    smuggle ～を密輸する

(1) Choose the appropriate summary for Paragraph 2-5. (2 点×4)

|             |    |
|-------------|----|
| Paragraph 2 | 61 |
| Paragraph 3 | 62 |
| Paragraph 4 | 63 |
| Paragraph 5 | 64 |

① Irene Sendler's story remained largely unnoticed until four high school students from Uniontown, Kansas brought it to light.

② She smuggled the children out using various means and placed them with Polish families, recording their names in a jar for future reference.

③ She was captured and tortured by the Nazis, refusing to reveal the hidden lists. A member of the Polish underground bribed a guard to release her, and she continued to rescue another 2,100 children while in hiding.

④ Irene Sendler persuaded Jewish parents to place their children in her care to save them from the Warsaw Ghetto.

(2) Which of the following title would be the most appropriate for this passage? 65 (2 点)

① "The Warsaw Ghetto: Life and Struggles During the Holocaust"

② "The Role of Żegota in World War II"

③ "High School Students Discover Historical Secrets"

④ "Irene Sendler: The Heroine Who Saved Thousands of Jewish Children During the Holocaust"

■図4：送付した定期試験問題の大問5 A(上智大学, 2010を参考に改変)

44

## 4.2.2 評価基準(ルーブリック)の適切さと明確さ

### 4.2.2.1 スピーキング

スピーキングの評価基準(ルーブリック)については、いくつかの観点から改善の余地が指摘された。フィードバックの対象となったスピーキングテストのルーブリックを表6に示す。対象となったルーブリックは言語的側面や表現内容に基づく五つの分析的評価観点「Accuracy」「Delivery」「Eye-contact」「Structure」「Content」に、「Accuracy」「Content」の観点は4レベル(S～C),「Delivery」「Eye-contact」「Structure」の観点は3レベル(A～C)の設計になっている。

まず、「Accuracy」のS評価にはイントネーションや発音の正確さも含まれているが、これはA～Cで評価している語彙・文法の誤りとは異なる観点であるため、A～Cの記述にもイントネーション等の評価観点を含める必要があるのではないかとの指摘があった。また、各ランクにおける記述文の表現と評価内容の一貫性にも注意が必要であり、例えば「語彙表現や文法の誤り」と「語彙や文法の誤り」など表記の揺れが見られるほか、誤りの頻度に加えて意味の伝わりやすさが評価に含まれている点も整理が求められる。さらに、「Eye-contact」のA評価における「プロフェッショナルに見える」という表現は曖昧であり、スピーキングの場面設定(高校生が中学生に動物について説明する場面)との整合性にも疑問が残るとの指摘があった。この点については、「プロフェッショナルに見える」といった抽象的な表現ではなく、「聞き手と適切にアイコンタクトをとり、自然で自信のある印象を与えている」など、観察可能で具体的な行動を基準とした表現に修正することが考えられる。

これらの評価観点の記述文の内容に加えて、ルーブリック全体として「Accuracy」「Delivery」「Eye-contact」「Structure」「Content」の5観点到問題はないものの、筆記テストと同様に「知識・技能」「思考・判断・表現」といった学習指導要領に基づく評価観点と対応づける必要があるのではないかという指摘もあった。国立教育政策研究所(2021)の参考資料は、「書くこと」の評価において、「知識・技能」を問う問題は主に英語使用の適切さ、「思考・判断・表現」を問う問題は主に表現内容の適切さで評価すると整理しており、「話すこと」の指導においても、「表現内容の適切さ」という内容面と、英語使用の適切さという言語面からの指導を行う」との記述から、同様の解釈が可能であることが考えられる。また、山本ら(2021)は、開発したスピーキングテストの評価基準における「正確さ」は「知識・技能」に、「内容」、「応答」、「質問」、「提案」は「思考・判断・表現」「主体的に学習に取り組む態度」に対応づけた。これらを踏まえると、本調査における上記5観点的な分析的な評価観点については、英語使用の適切さに関連する「Accuracy」「Delivery」「Eye-contact」を「知識・技能」へ、表現内容の適切さに関連する「Structure」「Content」を「思考・判断・表現」「主体的に

■表6: スピーキングテストのルーブリック

|   | Accuracy                                            | Delivery                                                | Eye-contact                  | Structure                                    | Content                                 |
|---|-----------------------------------------------------|---------------------------------------------------------|------------------------------|----------------------------------------------|-----------------------------------------|
| S | (Aに加えて)<br>語や句、文におけるイントネーションなどの発音の特徴をとらえ、正しく発音している。 |                                                         |                              |                                              | (Aに加えて)<br>具体例を効果的に示していて、情報量が多い。        |
| A | ・語彙表現や文法の誤りがほとんどない。<br>・聞き手に負荷を与えない。                | 以下の条件を全て満たしている。<br>・相手に十分伝わる声で話している。<br>・制限時間内で話し続けている。 | 十分なアイコンタクトがあり、プロフェッショナルに見える。 | 「生態」→「特徴」→「リスクや危機」→「解決策、メッセージ」という構成で話を伝えている。 | ・重要な情報を漏れなく伝えられている。<br>・話の流れが明確で一貫性がある。 |
| B | ・語彙や文法の誤りが時々あるが、意味は伝わる。<br>・聞き手に負荷を与えることがある。        | 上記条件を一つ満たしている。                                          | スピーチの半分以上でアイコンタクトが取れる。       | 上記の構成に多少のずれがある。                              | ・重要な情報に多少の漏れがある。<br>・話の順序に多少の前後がある。     |
| C | Bに満たない。                                             | 上記条件を一つも満たしていない。                                        | Bに満たない。                      | 上記の構成にずれが多く、何をまとまりで話しているか伝わらない。              | ・重要な情報に多くの漏れがある。<br>・話の順序に一貫性がない。       |

学習に取り組む態度」へ対応づけることが考えられる。その結果、言語的側面や表現内容に基づく分析的評価観点を設定したルーブリックを活用しつつも、ルーブリックの評価観点と観点別評価の評価観点の対応づけを明示することで、学習指導要領の目標に基づいて測定したい能力を測るテストとしての妥当性がより高まることが考えられる。一方で、観点別評価との対応を前提とした「知識・技能」「思考・判断・表現」「主体的に学習に取り組む態度」の3観点をそのまま用いたルーブリック設計に疑義を呈する見解もあり、今後の検討においてこうした批判的見解も考慮する必要がある(駒井, 2024)。

4.2.2.2 ライティング

ライティングテストのルーブリックについては、評価観点の表記に一貫性があり、基準も明確である点は高く評価された。表7に示すとおり、対象となったルーブリックは「内容」「構成」「正確さ」の言語的側面や表現内容に基づく三つの分析的評価観点から成り立っており、記述は簡潔かつ分かりやすいものであると評価された。一方で、スピーキングのルーブリックと同様に、設定された評価観点それ自体には問題はないものの、筆記テストにおける他の技能(リーディングやリスニング)では「知識・技能」「思考・判断・表現」など、学習指導要領に基づく観点別評価が行われていることを踏まえると、ライティングにおいてもそれらの観点と対応づけたルーブリック設計が望ましいのではないかという指摘があった。

例えば、英語使用の適切さに関連する「正確さ」の観点は「知識・技能」に対応し、表現内容の適切さに関連する「構成」「内容」の観点は「思考・判断・表現」に対応しうると考えられる。こうした言語的側面や表現内容に基づく分析的評価観点と学習指導要領における観点別評価の3観点との関係性を明示することで、教室内のパフォーマンス評価と制度的な評価の枠組みとの整合性がより高まることが期待される。

さらに、複数パラグラフの英文を書く課題において、「内容伝達に支障をきたす誤り」が複数出現する可能性が高いという懸念も示された。例えば、ある生徒が序論—本論—結論の3パラグラフを構成した場合、一つのパラグラフ内で主語・動詞の欠落や語順の誤りなどが複数含まれると、それだけでC評価に分類される可能性が高くなる。その結果、多くの受験者がC評価になってしまうとパフォーマンス評価の妥当性を下げる要因にもなり得るため、ルーブリックの記述内容やA～Cの基準設定において慎重な検討が求められる。

表7: ライティングテストのルーブリック

|    | 内容                                                 | 構成                                                     | 正確さ                                           |
|----|----------------------------------------------------|--------------------------------------------------------|-----------------------------------------------|
| A  | 内容に一貫性があり、例・理由などに説得力がある。                           | 序論—本論—結論の構成が確立し、さらに結論に工夫が見られる。                         | 内容伝達に支障をきたす誤りがない。                             |
| B  | 内容に一貫性がある。                                         | 序論—本論—結論の構成が確立している。                                    | 内容伝達に支障をきたす誤りが一つある。                           |
| C  | 内容に一貫性がない。                                         | 序論—本論—結論の構成が確立していない。                                   | 内容伝達に支障をきたす誤りが二つ以上ある。                         |
| 備考 | 内容の面白さではなく、与えられたトピックに関して首尾一貫しているか、付加情報があるかなどを判断する。 | 「結論に工夫」とは、単に序論を繰り返すのではなく、言い換えたり、判断・感想などを加えたりしていることとする。 | S+V構造の欠如、語順の誤り、時制の誤りを含むものなどに加えて、意味が通じないものも含む。 |

4.2.3 試験全体の妥当性や改善すべき点について

本調査においては、学習内容の定着度を測定する到達度テストとして、① Can-do リスト・指導と評価の計画とテスト内容が一致しているか、②学習指導要領に基づいた観点別評価(知識・技能、思考・判断・表現)が適切に行われているか、の観点で検討された。全体としては、定期試験はリーディング及びリスニ

ングを中心に構成されており、文法・語法・語彙といった言語知識も含まれている点から、ある程度は授業内外の学習内容に即した出題であったとの評価が得られた一方で、学習内容との関連性や観点別評価観点との整合性について複数の課題が指摘された。

大問1では、小問Aはディクテーションの形式を取っているが、単元の指導と評価の計画においてディクテーション活動が明記されておらず、学習内容との関連性がやや不明瞭であるとされた。したがって、ディクテーションを評価に用いる場合には、計画段階でその活動を明記し、授業内外での学習を経た上で試験に組み込むことが妥当であると考えられる。大問1の小問B及びCについては、「待ち合わせの場所について話しています」といった簡潔な場面が提示されてはいるものの、場面や目的、聞き手の立場などの情報が不十分であり、「思考・判断・表現」の観点からの評価には適していない可能性があるとの指摘があった。図5に大問1の設問内容を示す。国立教育政策研究所(2021, p. 64)による「聞くこと」の「思考・判断・表現」を評価する問題例においては、例えばKasumi and her classmate Michael are talking in their classroom. Listen to the dialogue and choose the best option from among the four choices. You will be able to listen to the dialogue only once.といった場面設定が導入されている。しかし、渡邊ら(2025)はこのような例について、場面のみが記述されており、目的や状況が明示されていない点を問題視している。さらに、大学入学共通テスト英語(リスニング)においても本調査の定期試験と同程度の場面設定が多く見受けられる他、場面設定の提示のない問題項目もある(大学入試センター, 2025b)。そのため、目的や状況をどの程度まで場面設定として明示すべきかについては、今後さらなる検討が必要であるが、学習指導要領に基づいて生徒の思考力・判断力・表現力を測る問題を作成するためには、より明確な場面設定を導入することは改善点の一つとして挙げられる。

大問4のA(1)に関しては、内容理解ではなく英文のつながりに焦点を当てた設問となっており、「思考・判断・表現」の観点で評価するには不適切ではないかとの指摘があった。大問4のAの設問内容は図6に示す通りであり、(1)では英文中に設けられた空所に対し、前後の文脈を手がかりに適切な語句を選択させる形式である。国立教育政策研究所(2021, p. 38)は、「読むこと」における「思考・判断・表現」の評価規準として、「コミュニケーションを行う目的や場面、状況などに応じて、日常的な話題や社会的な話題について、必要な情報を読み取り、書き手の意図や概要、要点を捉えている」とこととしている。この規準に照らすと、空所補充のように語彙や接続表現の知識に依存する設問は、書き手の意図や概要・要点の把握には直接関与せず、「知識・技能」の観点に分類するのが適切であると考えられる。一方、渡邊ら(2025)は、「読むこと」の「知識・技能」と「思考・判断・表現」を区別する際に、「内容を捉える技能」と「概要・要点を捉えている状況」を明確に分けた問題作りの難しさを指摘している。実際にこのようなテスト項目と評価観点の分類との齟齬は、村越・江原(2015)においても指摘されており、平成29年・30年告示学習指導要領下においても同様の問題が依然として存在していることが示唆される。

次に、技能ごとの配点の割合については、リスニング20%、リーディング40%、文法・語法20%、語彙20%という構成であった。この配分に対しては、英語4技能の育成を目的とする英語コミュニケーションⅢの評価としては、文法・語法・語彙に対する比重がやや高く、一方でリスニングが相対的に少ないという課題が指摘された。この配分がCan-doリストや単元の指導と評価の計画と整合的であるかについては、改めて検証する必要がある。実際に、勤務校のCan-doリスト及び単元の指導と評価の計画においては、英語4技能をバランスよく指導・評価することが示されており、リスニングも他技能と同等の比重で配点されることが理想である。なお、本研究では、定期試験及びパフォーマンステストが年間成績に占める割合や、両者を合わせた全体の評価比率を正確に算出することは困難であった。そのため、ここで示す配分は定期試験内での技能比率に限られる点に留意する必要がある。

また、観点別評価における配点の割合は、「知識・技能」48%、「思考・判断・表現」52%と一見バランスが取れているように見えるが、前述のように観点のラベリングに齟齬が含まれている可能性がある。そのため、各設問がどの評価観点に適切に分類されているかを再確認した上で、実際にバランスの取れた評価が行われているかを見直す必要があるとの指摘があった。

1【リスニング】 知識・技能 [8] 思考・判断・表現 [12]

【A】〈記述問題〉以下の英文はこれから読まれる英文のスク립トです。英文を聞き、空所(あへく)に読まれた語または数字を書き入れ、解答用紙にその答えを書きなさい。1 回避します。(知技 1 点×8)

(※解答欄は解答用紙の裏面にあります。)

You might think that the traditional French ham and butter baguette sandwich would be the (あ) popular fast food in France. But think again. A recent study shows that sales of American-style hamburgers have (い) the sales of these sandwiches. According to a French restaurant marketing company, about 1.2 billion ham and butter baguette sandwiches were sold in France in 2017, while 200 (う) more burgers were sold than sandwiches during the same period. Sales of both baguette sandwiches and hamburgers are on the (え), the study says. It notes that baguette sandwiches saw 1.3 percent growth in 2017, while burger sales recorded a 9 percent (お). Curiously, in France, the number of hamburgers sold by a major fast food restaurant chain has not varied for years. What has changed, however, is the (か) number of restaurants putting burgers on their menus. In fact, (き) percent of restaurants surveyed in France included burgers on their menus in 2017. The rise of the American-style burger was noted in a newspaper story which reported that many famous restaurants in Paris (く) burgers to their menus for the first time 10 years ago.

【B】〈マークシート問題〉対話とそれについての問いを聞き、答えとして最も適切なものを一つずつ選びなさい。2 回避します。(思判表 2 点×3)

問1 待ち合わせ場所について話をしています。1



問2 昨晚のパーティについて話をしています。2

- ① Bored.
- ② Concerned.
- ③ Funny.
- ④ Surprised.

問3 夫婦がお店で話をしています。3

- ① \$100
- ② \$120
- ③ \$140
- ④ \$160

【C】〈マークシート問題〉対話を聞き、問いの答えとして最も適切なものを一つずつ選びなさい。1 回避します。(思判表 2 点×3)

問1 男性が女性に今朝の状況を説明しています。4

- How did the man arrive?
- ① By airplane.
  - ② By car.
  - ③ On the 9:30 train.
  - ④ On the 10:00 train.

問2 ホームパーティについて話をしています。5

- Where is the man's house?
- ① Across from the high school.
  - ② Around the corner from the women's house.
  - ③ In the neighborhood of the city office.
  - ④ Next to the convenience store.

問3 映画について話をしています。6

- Which of the following is true about the women?
- ① They didn't watch the movie in the afternoon.
  - ② They each spent \$10 to enter the movie.
  - ③ They enjoyed drinking instead of watching the movie.
  - ④ They paid as much for drinks for watching the movie.

■図5: 対象となった定期試験問題の大問1 (A~C)

## 4【FLEX】知識・技能〈マークシート問題〉思考・判断・表現 [20]

[A] You are taking a lecture about optical illusions and reading the following text. Read the text and answer the questions below.

Take a look at the object in the illustration below. What does it look like to you? To some people this might appear to be a rabbit facing to the right. (47), it may seem like a duck looking in the opposite direction to others. This illustration is an example of what is called an "optical illusion", (48) it shows how our brains can interpret the same image in different ways. Illusions such as these have a variety of uses and are often seen around us in our daily lives.

To start, these types of illusions have traditionally been incorporated into various structural and architectural designs. For example, the *karesansui* (dry landscape) garden at the Zen temple Ryōanji in Kyoto illustrates this concept well. At first glance, its entire surface may appear perfectly flat, but if we look closer, we can see that it actually inclines towards the garden's south-east corner. (49), the earthen wall on the west-side is slightly higher on its northern end. These features combine to create an optical illusion and to our eyes the surface actually appears to be deeper than it truly is. This same effect can also be seen in the Ponzo Illusion. In the Ponzo Illusion, the uppermost horizontal line appears to be longer than the line just below it, but (50), the two lines are exactly the same length.

Second, if we look at the image below, we can see another example of an optical illusion. The interior of the house appears to have an extremely high ceiling though it is actually much lower than it looks. Our brains interpret the image in this way as a result of the Fick Illusion (Figure 2). In the Fick Illusion, the vertical line seems longer than the horizontal line. (51), they are exactly the same length. This effect can be very useful in architecture, as it can cause us to think a room is much more spacious than it actually is.

Third, another valuable technique which has been used successfully in recent times is the creation of what appears to be a three-dimensional object on a two-dimensional surface. (52), in one particular train station people are often unable to find an elevator because it is located in a blind spot, so a sign that appears three-dimensional has been painted on the floor in order to help them find it more easily. The sign quickly catches people's attention and has been extremely successful in helping reduce the number of people who have been become lost while trying to locate the elevator.

Last, they have been found to be useful in the field of transportation, in particular at intersections which don't have traffic lights. Crosswalk markers which appear to be three-dimensional have been painted on some roads in an effort to prevent traffic accidents from occurring. From the viewpoint of approaching drivers, it looks like they are about to hit large concrete blocks which are in the road. (53), these were created in hopes that the drivers will see them and slow down as they get nearer the crosswalk.

Optical illusions are just illusions. They cannot magically make a garden larger, nor can they truly make a ceiling higher. Over time, even if a crosswalk is painted to appear three-dimensional, it will eventually lose its effectiveness in preventing accidents as drivers who regularly see these markers will get used to them and come to understand that what they are seeing is just an illusion and they won't slow down. (54), even if we recognize such common limitations of these illusions, we should be able to find numerous other useful applications for optical illusions in society.

(1) Put an appropriate word or phrase in each blank in the text. (1点×8)

- |    |                    |               |                    |                 |
|----|--------------------|---------------|--------------------|-----------------|
| 47 | ①On the other hand | ②In fact      | ③As a result       | ④Besides        |
| 48 | ①conversely        | ②therefore    | ③for example       | ④in other words |
| 49 | ①However           | ②Besides      | ③On the other hand | ④For instance   |
| 50 | ①in conclusion     | ②in short     | ③in fact           | ④meanwhile      |
| 51 | ①In general        | ②Consequently | ③Moreover          | ④However        |
| 52 | ①Additionally      | ②For instance | ③Therefore         | ④As a result    |
| 53 | ①On the contrary   | ②In fact      | ③Thus              | ④Furthermore    |
| 54 | ①To sum up         | ②Nevertheless | ③In addition       | ④In other words |

(2) Which is the main point of the text? (2点) 55

- ① Optical illusions can create the illusion of depth in gardens and structural designs.  
 ② Optical illusions are adaptable tools used in various aspects of daily life, including architecture, navigation, and transportation.  
 ③ The effectiveness of optical illusions diminishes over time as people get accustomed to them.  
 ④ The illustration of the rabbit and duck is a classic example of how optical illusions work.

パフォーマンステストについては、全体として授業内容との関連性は高く評価された。例えば、スピーキングテスト(図2参照)では、教科書に登場する動物に関するトピックを基に「中学生に向けて動物について説明する」といったタスクが設定されており、学習した語彙や表現を活用できる設計となっていた。このように学習内容との整合性は見られた一方で、タスクと単元の指導と評価の計画との整合性には課題が見られた。具体的には、評価観点に「Eye-contact」や「Delivery」が含まれていたが、これらの技能については指導と評価の計画に記載がなかったため、計画上の学習活動との対応が不十分であるとの指摘があった。

ライティングテスト(図3参照)については、リーディング活動と組み合わせた技能統合型のタスクとして構成されており、与えられた英文資料を読み取った上で、自分の意見を英語で記述するという形式が取られていた。この点は、実際の言語使用場面に近いタスク設計であるとして肯定的に評価された。しかしながら、評価の中心が「思考・判断・表現」に置かれているにもかかわらず、「誰に向けて、どのような状況で書くのか」といったコミュニケーションの目的や場面、状況設定が明示されていなかったことから、評価基準の前提条件が不明瞭であるとの指摘があった。国立教育政策研究所(2021, p. 76)の参考資料の例においては、「下記は海外交流校の生徒が送ってきた質問です。質問を読み、どちらがより効果的だと思うか英語で書きなさい。」という指示文が提示されており、ライティングに際して誰に向けて書くべきかが明確になっている。本調査で対象となった「ソーシャルメディアの利点や欠点についての意見を書く」というタスクであれば、「海外交流校の高校生との議論に向けて」など、ライティングの目的や想定読者を明示することで、学習者の表現の方向性が明確になり、評価の妥当性も高まると考えられる。

これらのことから、パフォーマンステストの設計においては、学習活動との一貫性を確保しながら、課題文の条件設定をより明確化するとともに、Can-doリストとの対応関係やルーブリックの評価観点の構造を可視化することが、妥当性の向上につながるといえる。また、Can-doリストに基づく学習到達目標が、実際の指導内容や評価活動と一貫性をもって連動しているかを検証することも求められる。最後に、本フィードバックを踏まえた今後の改善点として、これまでに指摘された事項に加え、テスト仕様書の作成が重要であるとの指摘があった。これにより、各テスト項目が測定しようとする能力(construct)を明確に定義することができ、テスト全体の妥当性の向上に資することが期待される。

### 4.3 波及効果推論に係る生徒質問紙調査

本調査では、定期試験後に勤務校の高校1年生から3年生までの生徒計1,103名を対象に、英語コミュニケーションの定期試験に関する質問紙調査を実施し、924名から回答を得た(回収率83.8%)。学年別の回答者数は、1年生370名、2年生300名、3年生254名であり、全学年から概ね均等に回答が集まった。各学年の英語コミュニケーションの定期試験は、概ね教科書を題材としたリーディング問題、文法・語法問題、さらには大学入試問題や外部英語試験を基にしたリーディング問題で構成されている。なお、リスニング問題は2年生の試験にのみ含まれており、1年生及び3年生については、定期試験とは別に、授業内でリスニング試験が実施されているため、本調査で扱った定期試験には含まれていない。また、生徒が回答した定期試験の内容は、本研究において専門家レビューの対象となった定期試験問題とは異なるものである。

Allen et al.(2024)の質問紙調査では、英語圏での滞在経験を有する生徒は全体の25.9%、将来的に英語圏での生活や留学を希望している生徒は68.5%であった。これに対し、本調査では滞在経験を有する生徒は11.6%、留学等を希望している生徒は50.2%にとどまった。これらの結果から、本研究の対象生徒は一定の関心を示してはいるものの、英語圏での経験及び将来的な関心のいずれにおいても、先行研究の対象生徒に比べて相対的に低い関心を示していることが分かった。生徒質問紙調査の質問項目及び調査結果並びに選択式回答項目の選択肢及び回答結果を資料6,7に示す。

本調査における授業に関する設問では、Q5「先生が話し、生徒が聞く時間が多い」( $M = 3.5$ )、Q6「先生は主に英語を話す」( $M = 3.9$ )といった項目には6段階中の中間値(3.5)と同程度及びやや高めの数値を示した。本調査では3と4のラベルが設定されていないため厳密な解釈は難しいが、平均値としては中程度及びや

や高めといえる。一方で、Q7「生徒が先生と話す際に英語を使う」( $M = 2.9$ )、Q8「生徒同士のやりとりで英語を使う」( $M = 3.2$ )といった生徒のアウトプットに関する項目は中間値よりも低い傾向を示した。ただし、Q9「即興で英語を話すことが求められることが多い」という項目については( $M = 3.7$ )と中間値よりもやや高めだった。全体としては教師によるインプットは比較的豊富である一方で、生徒の英語によるアウトプットの機会は限定的であることがうかがえる。これらの結果は、今後の授業改善において、生徒の発話を促進する授業設計の必要性を示唆するものである。一方、Allen et al. (2024)の同様の設問においては、Q5「先生が話し、生徒が聞く時間が多い」( $M = 2.4$ )と低く評価される一方、Q6「先生は主に英語を話す」( $M = 5.3$ )、Q7「生徒が先生と英語で話す」( $M = 4.2$ )、Q8「生徒同士で英語を使う」( $M = 4.2$ )、Q9「即興で英語を話すことが求められる」( $M = 5.1$ )と、アウトプット重視の授業実践がより積極的に行われている様子が示されている。この比較から、本研究の対象校では教師主導の授業が中心であるのに対し、Allen et al. の対象校では生徒の主体的な英語使用がより重視されていることがうかがえる。今後は、生徒の発話機会を体系的に増やすための授業デザインの検討が、よりバランスの取れた言語活動の促進に寄与すると考えられる。

Q10「英語コミュニケーションの授業の活動は主に何に焦点を当てていますか？(複数回答可)」という設問に対しては、「スピーキング」(58.9%)、「リスニング」(58.6%)、「リーディング」(56.9%)といった技能項目が語彙や文法などの知識項目より高い割合で選ばれた(表8)。学年別に見ると、リスニングは3年生で最も高い割合(63.4%)を示し、定期試験に含まれていないにもかかわらず、生徒は授業で重視されていると認識していた。一方で2年生は定期試験にリスニング問題が含まれているにもかかわらず、その割合は55.7%と低く、授業と試験設計の方針が必ずしも一致していないことが示唆された。特に3年生では、大学入学共通テストにおけるリスニング配点の増加や英検などの外部検定試験対策が授業内容に影響している可能性があるが、本調査はその因果関係を直接検証するものではなく、教師・生徒へのインタビューなど他の方法による裏付けが必要である。

一方、「スピーキング」に焦点を当てていると回答した割合は、1年生71.4%、2年生53.7%、3年生46.9%と学年が上がるにつれて低下した。これはAllen et al. (2024)が指摘するように、大学入試におけるスピーキング技能の軽視が授業内容や学習活動の変化に負の波及効果(negative washback)を与えている可能性がある。

■表8: Q10「英語コミュニケーションの授業の活動は主に何に焦点を当てていますか？」(複数回答可)

|    |      | 語彙    | 文法    | 発音    | リスニング | リーディング | スピーキング | ライティング |
|----|------|-------|-------|-------|-------|--------|--------|--------|
| 全体 | 回答件数 | 474   | 414   | 405   | 541   | 526    | 544    | 234    |
|    | 回答率  | 51.4% | 44.8% | 43.8% | 58.6% | 56.9%  | 58.9%  | 25.3%  |
| 1年 | 回答件数 | 200   | 152   | 211   | 213   | 214    | 264    | 73     |
|    | 回答率  | 54.1% | 41.1% | 57.0% | 57.6% | 57.8%  | 71.4%  | 19.7%  |
| 2年 | 回答件数 | 148   | 159   | 105   | 167   | 144    | 161    | 116    |
|    | 回答率  | 49.3% | 53.0% | 35.0% | 55.7% | 48.0%  | 53.7%  | 38.7%  |
| 3年 | 回答件数 | 127   | 103   | 89    | 161   | 168    | 119    | 45     |
|    | 回答率  | 50.0% | 40.6% | 35.0% | 63.4% | 66.1%  | 46.9%  | 17.7%  |

注:  $N=924$ 。複数回答を可としたため、回答率の合計は100を超える。

Q24「今回の定期試験のために学校外では主に何を中心に学習しましたか？」(複数回答可)への回答では、多くの生徒が語彙(59.4%)、リーディング(49.7%)、文法(41.6%)を重点的に学習しており、それ以外の領域は低い割合にとどまった(表9)。特にリスニングは、1年生(2.7%)や3年生(2.4%)で極めて低く、2年生では21.7%と相対的に高かったものの、依然として他の領域に比べて低水準であった。この結果は、2年生で

リスニング問題が定期試験に含まれていることが学習行動に一定の影響を与えている可能性を示す一方で、その波及効果は限定的であることも示唆している。その要因の一つとして、定期試験におけるリスニングの配点が比較的低く、学習者にとって優先度が下がってしまうことが考えられる。一方で、ライティングやスピーキングは定期試験ではなくパフォーマンステストで評価されるため、定期試験における波及効果はほとんど見られない。今後は、パフォーマンステストにおける波及効果についても検証する必要がある。

表9: Q24「今回の定期試験のために学校外では主に何を中心に学習しましたか?」(複数回答可)

|    |      | 語彙    | 文法    | 発音   | リスニング | リーディング | スピーキング | ライティング | ない    |
|----|------|-------|-------|------|-------|--------|--------|--------|-------|
| 全体 | 回答件数 | 549   | 384   | 40   | 81    | 459    | 56     | 102    | 74    |
|    | 回答率  | 59.4% | 41.6% | 4.3% | 8.8%  | 49.7%  | 6.1%   | 11.0%  | 8.0%  |
| 1年 | 回答件数 | 228   | 153   | 17   | 10    | 192    | 26     | 61     | 25    |
|    | 回答率  | 61.6% | 41.4% | 4.6% | 2.7%  | 51.9%  | 7.0%   | 16.5%  | 6.8%  |
| 2年 | 回答件数 | 157   | 140   | 18   | 65    | 154    | 19     | 33     | 21    |
|    | 回答率  | 52.3% | 46.7% | 6.0% | 21.7% | 51.3%  | 6.3%   | 11.0%  | 7.0%  |
| 3年 | 回答件数 | 164   | 91    | 5    | 6     | 113    | 11     | 8      | 28    |
|    | 回答率  | 64.6% | 35.8% | 2.0% | 2.4%  | 44.5%  | 4.3%   | 3.1%   | 11.0% |

注. N=924. 複数回答を可としたため、回答率の合計は100を超える。

スコアレポート(得点個表)に関する設問では、Q30「自身の強みや弱点を理解するのに役立つ」、Q31「内容が明確である」、Q32「解釈しやすい」(いずれもM = 4.6)といった肯定的な評価が得られた。学年によって多少の違いはあるものの、勤務校で用いられているスコアレポートは生徒の自己評価や学習の振り返りを促す有効な資料として機能していることが示唆される(図7参照)。一方で、小泉ら(2023)の研究では、スコアレポートを有効に活用するためには、単なる情報提供にとどまらず、解説動画や自己分析活動と組み合わせることで、学習への結び付きが強化されることが指摘されている。こうした知見を踏まえると、勤務校におけるスコアレポートの活用においても、どのように活用されているかを検討し、生徒のより自律的な学習支援へとつなげていく余地があると考えられる。

| 11組 1番 氏名 |    |     |    |     |    |      | 模範解答 |     |     |    |    |    |      |
|-----------|----|-----|----|-----|----|------|------|-----|-----|----|----|----|------|
| 番号        | 正答 | マーク | 判定 | 観点  | 得点 | 正答率  | 番号   | 正答  | マーク | 判定 | 観点 | 得点 | 正答率  |
| (1)       | 2  | 2   | ○  | 思判  | 2  | 54.2 | (28) | 3   | 3   | ○  | 知技 | 2  | 93.2 |
| (2)       | 1  | 1   | ○  | 思判  | 2  | 76.3 | (29) | 1   | 1   | ○  | 知技 | 2  | 93.2 |
| (3)       | 4  | 4   | ○  | 思判  | 2  | 47.6 | (30) | 4   | 4   | ○  | 知技 | 2  | 67.8 |
| (4)       | 1  | 1   | ○  | 思判  | 2  | 9.1  | (31) | 1   | 1   | ○  | 知技 | 2  | 13.1 |
| (5)       | 3  | 3   | ○  | 思判  | 2  | 40.3 | (32) | 5   | 5   | ○  | 知技 | 2  | 32   |
| (6)       | 3  | 3   | ○  | 思判  | 2  | 51.9 | (33) | 4   | 4   | ○  | 知技 | 2  | 44.1 |
| (7)       | 3  | 3   | ○  | 思判  | 2  | 37.5 | (34) | 1   | 1   | ○  | 知技 | 2  | 44.1 |
| (8)       | 4  | 4   | ○  | 思判  | 2  | 62   | (35) | 2   | 2   | ○  | 知技 | 2  | 29   |
| (9)       | 1  | 1   | ○  | 思判  | 2  | 32   | (36) | 3   | 3   | ○  | 知技 | 2  | 29   |
| (10)      | 1  | 1   | ○  | 思判  | 2  | 54.2 | (37) | 5   | 5   | ○  | 知技 | 2  | 32   |
| (11)      | 3  | 3   | ○  | 知技  | 2  | 71   | (38) | 3   | 3   | ○  | 知技 | 2  | 32   |
| (12)      | 2  | 2   | ○  | 知技  | 2  | 90.2 | (39) | 5   | 5   | ○  | 知技 | 2  | 32   |
| (13)      | 4  | 4   | ○  | 知技  | 2  | 66.7 | (40) | 6   | 6   | ○  | 知技 | 2  | 59.7 |
| (14)      | 4  | 4   | ○  | 知技  | 1  | 64   | (41) | 4   | 4   | ○  | 知技 | 2  | 61.5 |
| (15)      | 4  | 4   | ○  | 知技  | 1  | 40.3 | (42) | 1   | 1   | ○  | 知技 | 2  | 38.3 |
| (16)      | 2  | 2   | ○  | 知技  | 1  | 67.8 | (43) | 3   | 3   | ○  | 知技 | 2  | 56.9 |
| (17)      | 2  | 2   | ○  | 知技  | 1  | 84.4 | (44) | 2   | 2   | ○  | 知技 | 2  | 48.4 |
| (18)      | 2  | 2   | ○  | 知技  | 1  | 67   | (45) | 2   | 2   | ○  | 知技 | 2  | 66.5 |
| (19)      | 1  | 1   | ○  | 知技  | 1  | 85.9 | (46) | 4   | 4   | ○  | 知技 | 2  | 78.1 |
| (20)      | 2  | 2   | ○  | 知技  | 1  | 79.8 | (47) | 5   | 5   | ○  | 思判 | 2  | 31.5 |
| (21)      | 3  | 3   | ○  | 知技  | 1  | 82.6 | (48) | 2   | 2   | ○  | 思判 | 2  | 61.2 |
| (22)      | 3  | 3   | ○  | 知技  | 1  | 79.3 | (49) | 3   | 3   | ○  | 思判 | 2  | 68.3 |
| (23)      | 2  | 2   | ○  | 知技  | 1  | 61.2 | (50) | 1   | 1   | ○  | 思判 | 2  | 65.7 |
| (24)      | 4  | 4   | ○  | 知技  | 2  | 22.4 | (51) | 2   | 2   | ○  | 思判 | 2  | 42.3 |
| (25)      | 3  | 3   | ○  | 知技  | 2  | 68.5 | (52) | 3   | 3   | ○  | 思判 | 2  | 55.7 |
| (26)      | 2  | 2   | ○  | 知技  | 2  | 68.5 | (53) | 1   | 1   | ○  | 思判 | 2  | 19.9 |
| (27)      | 4  | 4   | ○  | 知技  | 2  | 68.5 | (54) | 2   | 2   | ○  | 思判 | 2  | 32.5 |
|           |    |     |    |     |    |      |      |     |     |    |    |    |      |
| マ         |    | 知技  |    | 46  |    | /    |      | 46  |     |    |    |    |      |
| イ         |    | 思判  |    | 36  |    | /    |      | 36  |     |    |    |    |      |
| タ         |    | 知技  |    | 10  |    | /    |      | 10  |     |    |    |    |      |
| 記         |    | 思判  |    | 8   |    | /    |      | 8   |     |    |    |    |      |
| 述         |    | 知技  |    | 56  |    | /    |      | 56  |     |    |    |    |      |
| 小         |    | 思判  |    | 44  |    | /    |      | 44  |     |    |    |    |      |
| 計         |    | 合計  |    | 100 |    | /    |      | 100 |     |    |    |    |      |

図7: 勤務校で用いられるスコアレポートのサンプル

注: 正答率は学年の正答率を示す。

## 5 考察

### 5.1 教師質問紙調査と専門家レビューを通じた総合的考察(RQ1及びRQ2)

本研究では、教室内英語テストの妥当性を論証に基づく妥当性検証の領域定義推論の観点から検討するため、現職教師に対する質問紙調査と、言語テストの専門家による文書レビューを実施した。これら二つの調査は、それぞれ異なる視点から現行のテスト実践における妥当性の課題と改善の方向性を浮き彫りにしている。

まず、教師質問紙調査の結果から、多くの教師が教科書との整合性や授業内容との連動を強く意識して定期試験やパフォーマンステストを設計している一方で、学習指導要領や単元の指導と評価の計画、Can-doリストとの接続に関しては、教師間での理解の不十分さやばらつきが見られた。しかしながら、本来教室での評価は、学習指導要領に示された目標に基づいて、Can-doリストや単元ごとの指導と評価の計画が作成され、それらと一貫性をもった形で設計・実施されるべきである。したがって、教師は教科書や授業内容のみならず、TLU領域を定義する制度的文書との整合性を含めた、全体の連動性を意識する必要がある。このような視点に基づく教師間の認識の共有と自治体などの支援が、教室評価の妥当性を高める上で不可欠であると考えられる。さらに、観点別評価における評価規準の理解に教師間でばらつきが見られた。これは、観点別評価を運用する際の測定対象(構成概念)や評価観点に関する理論的理解と共通認識が十分に共有されていないことを示唆している。このように、評価の設計が表面的な整合性にとどまり、構成概念や能力記述に基づいた評価規準の理解が曖昧なまま運用される場合、評価の信頼性及び妥当性を損なうリスクがある。

これらの傾向は、専門家レビューの結果とも一致する部分がある。専門家からは、定期試験やパフォーマンステストにおいて、学習内容との整合性は概ね良好であるとの評価がなされた一方で、ルーブリックの言語的側面や表現内容に基づく分析的評価観点と学習指導要領に基づく評価観点(知識・技能、思考・判断・表現)との対応関係が不明確であること、TLU領域における測定すべき能力の明示が一部不十分であることが課題として指摘された。このことは、現場教師が観点別評価における構成概念を統一的に捉えることの難しさを示しており、自治体などによる研修を含めた実践・理論の両面からの支援が必要であることを示唆している。

以上の結果から、現行の教室内英語テストは、授業実践に基づいた設計という点では一定の妥当性を有するものの、TLU領域が定義されているとされる学習指導要領やCan-doリスト、単元の指導と評価の計画などの制度的な文書との整合性について、さらに観点別評価における測定すべき構成概念についての教師間の共通理解について課題が見られることが明らかとなった。したがって、本結果は領域定義推論における前提である「高等学校における教室内英語テストにおいてTLU領域は明確に定義されており、テストのタスクがTLU領域における重要なスキルを適切にカバーしている」を一部支持するものの、上記のように改善すべき課題が見出された。これらの課題は、今後の教室評価の改善に向けた重要な論点であると考えられる。

### 5.2 生徒質問紙調査の考察(RQ3)

本調査の結果から、定期試験が生徒の学習に対して一定の波及効果を及ぼしている可能性が示唆された。まず、試験準備では、多くの生徒が語彙(59.4%)、リーディング(49.7%)、文法(41.6%)といった言語知識やリーディングに重点を置いており、試験後に「自信がついた」項目でもリーディング(29.1%)、語彙(25.0%)、文法(24.2%)が上位を占めた。さらに、試験後に「さらに練習したい」と回答した項目もリーディング(58.0%)、語彙(57.7%)が高く、定期試験を契機とした継続的学習意欲の喚起が示唆される。一方で、2年生にのみ定

期試験にリスニング問題が含まれていたものの、他の領域と比べて低水準であった。対象が異なる試験であるが、専門家によるフィードバックにおいて定期試験のリスニングの配点が小さいという指摘も踏まえると、リスニングは授業で一定程度扱われているにもかかわらず、定期試験での出題の有無や配点が他の領域よりも少ないことが、学習者の優先度を下げ、プラスの波及効果を制限している可能性も考えられる。大学入学共通テストにおいてリーディングとリスニングの配点が同じであることも踏まえると、今後は教室内評価におけるリスニングの比率を高める必要があるかもしれない。また、ライティングやスピーキングは定期試験ではなくパフォーマンステストで評価されることを前提としているため、定期試験における波及効果はほとんど見られない。今後は、パフォーマンステストの波及効果についても検証する必要がある。

また、「試験の内容が授業と似ていた」「授業が試験準備に役立った」といった項目は平均4前後と中間値よりもやや高い評価を得ており、授業と試験内容の整合性がある程度確保されていることが分かる。これは、授業で学んだ内容が試験準備へと直接つながるというプラスの波及効果がある程度現れていると考えることができる。しかし、6段階評価において満点に近い値ではないことから、その効果は十分に最大化されているとは言い難く、改善の余地がある。さらに、授業の内容がそのまま試験に出題され、試験準備が単なる暗記になってしまう可能性があり、この点については別途インタビュー調査などで検証が必要である。また、専門家によるフィードバック調査における、定期試験の問題項目について目的・場面・状況の設定が不十分であるという指摘を踏まえると、日頃の授業から目的・場面・状況を明確に設定した言語活動を行い、それを評価タスクに反映させることが、授業と試験の一貫性をより高め、プラスの波及効果をより引き出すことにつながる可能性がある。

スコアレポート(得点個表)については、「自身の強みや弱点を理解するのに役立つ」( $M = 4.3$ )、「内容が明確である」「解釈しやすい」とともに( $M = 4.6$ )といった肯定的評価が得られ、生徒の自己評価や学習の振り返りを促す有効な資料として機能していることが示唆された。しかし、6段階評価において満点に近い値ではなく、依然として改善の余地がある。改善の方向性としては、小泉ら(2023)が指摘するように、単なる情報提供にとどまらず、解説動画や自己分析ワークシートとの組み合わせ、授業内での振り返り活動や教師による具体的なフィードバックの導入などが考えられる。これにより、スコアレポートの活用がより能動的な学習行動につながりやすくなると期待される。

以上より、本研究の対象となった定期試験は、特にリーディングや語彙・文法といった領域でプラスの波及効果を示す兆候がある一方で、リスニングへの波及効果は限定的であり、スコアレポート活用にも改善の余地がある。これらは波及効果推論における前提「高等学校における教室内英語テストは生徒に有益な波及効果を与えている」を一部支持しつつも、一部反論する結果となっていることが考えられるため、定期試験やスコアレポートの活用についてさらなる改善が求められる。しかしながら、パフォーマンステストの波及効果の検証など、本研究では検討できていない要素が多数存在することに留意が必要である。また、本研究で得られた結果は、生徒質問紙調査という自己報告データのみに基づくものであり、実際の授業行動や学習成果を直接観察したものではない。そのため、定期試験やパフォーマンステストが学習に及ぼす波及効果の全体像を把握するには限界がある。今後は、授業観察や生徒・教師へのインタビュー、学習成果の縦断的測定など、多角的な方法を組み合わせることで、より多面的で信頼性の高い検証を行う必要がある。

## 6 結論

本研究では、高等学校における教室内英語テストの妥当性を、Chapelle and Voss(2021)の論証に基づく妥当性検証の枠組みにおける「領域定義推論(Domain Definition Inference)」及び「波及効果推論(Consequence Implication Inference)」の観点から検討した。現職教師に対する質問紙調査、言語テストの専門家による文書レビュー、生徒に対する質問紙調査を通して、教室内テストの設計・運用に関する

現状と課題を明らかにした。

まず、RQ1「教室内英語テストにおいて目標言語使用領域は明確に定義されているか」については、多くの教師が教科書や授業内容との整合性を意識して評価を設計している一方で、学習指導要領、Can-do リスト、単元の指導と評価の計画といった制度的な文書との接続に対する理解や実践にはばらつきが見られ、TLU 領域の定義が必ずしも明確に共有・運用されているとは言えない点が確認され、領域定義推論の前提を一部支持するものの、改善が必要であることが分かった。

次に、RQ2「テストのタスクが目標言語使用領域の重要なスキルを適切にカバーしているか」については、学習内容との整合性は一定程度確認されたものの、タスクの設計において測定すべき構成概念の明示が十分でない場合や、教師によって分類された評価観点のラベリングに齟齬があること、さらにはルーブリックの評価観点と学習指導要領に基づく評価観点との対応関係が不明確であるといった課題が指摘された。これらは、領域定義推論の前提を一部支持するものの、評価設計における理論的な裏付けの不足や教師間の共通認識について今後の改善の必要性を示唆している。

また、RQ3「教室内英語テストは生徒にどのような波及効果を与えているか」については、特にリーディングや語彙・文法といった領域においてはプラスの波及効果の兆候が確認された一方で、リスニングについての波及効果は限定的であり、プラスの波及効果を与えていない兆候も見られた。これらは、波及効果推論の前提を一部支持しつつも、一部反論する結果となっていることが考えられるため、定期試験における配点の割合のバランスなどの改善が求められる。また、スピーキングやライティングといったアウトプット技能については、別途実施されるパフォーマンステストの波及効果の検討が求められる。

本研究で扱った教師質問紙調査・生徒質問紙調査・専門家レビューはいずれも筆者の勤務校における実施に限定されており、結果の一般化には限界がある。また、本研究では質問紙調査と文書レビューを主な方法として採用したが、教師や生徒へのインタビュー、授業観察、テスト結果の得点分析など、より多角的なアプローチを取り入れることで、より深い洞察や裏付けが得られる。したがって、今後は複数校を対象とした調査の実施や混合研究法による検証を通じて、教室内英語テストの妥当性に関する理解をさらに深めていく必要がある。

本研究は、高等学校における教室内英語テストを対象に、論証に基づく妥当性検証の枠組みを適用した、国内では先駆的な試みである。これまでの妥当性検証は主に大規模標準化テストに焦点を当てたものが多く、日常的に実施される教室内テストにこの枠組みを応用した研究は極めて限られていた。こうした背景のもと、本研究は、実際の教育現場で行われる定期試験やパフォーマンステストについて、領域定義推論及び波及効果推論という二つの視点から妥当性を検討し、評価実践に内在する課題を可視化するとともに、今後の改善に向けた示唆を提供するものである。今後は本研究を皮切りにして、さらに領域定義推論と波及効果推論の証拠を補強するとともに、得点化推論をはじめ、一般化推論、説明推論、外挿推論、利用推論の他の推論についても継続的に検討することが求められる。約10年ごとに改訂される学習指導要領に基づく学習評価については、本研究で採用した論証に基づく妥当性検証の枠組みなどの系統的な理論的枠組みを用いて、その妥当性を検討することが重要である。本研究はその端緒を示すものであり、今後、この枠組みを活用した検証が進むことで、より適切な指導と評価の実現につながることを期待される。

## 謝辞

はじめに、本研究を実施する機会を与えてくださった公益財団法人 日本英語検定協会の皆様、及び選考委員の皆様にご心より感謝申し上げます。特に、研究助言者である筑波大学の小泉利恵先生には研究実施から執筆にあたり貴重なご指導を頂戴いたしました。深く感謝申し上げます。また、本研究にあたって貴重なご意見をくださった琉球大学の深澤真先生をはじめ、英検研究助成についてご紹介くださり、日々の研究活動の中で多くのご助言をいただいている横浜国立大学の斉田智里先生、及び同大学院ゼミの皆様にも深く感謝申し上げます。最後に、本研究にご協力していただきました高校生並びに先生方に感謝申し上げます。なお、本報告書に関して、開示すべき利益相反関連事項はありません。

## 引用文献

- Allen, D., Yamamura, K., Meguro, S., & Nakamura, T. (2024). Washback from four-skills English exams in the Japanese secondary school context. *Language Assessment Quarterly*, 21, 473-505. <https://doi.org/10.1080/15434303.2024.2440895>
- 馬場哲生 (2023). 日本の中・高等学校の英語教育における観点別学習状況の評価: 歴史的変遷と今後の課題. 東京学芸大学紀要, 74, 467-481. <https://u-gakugei.repo.nii.ac.jp/records/55991>
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- Chapelle, C., Enright, M., & Jamieson, J. (2008). *Building a validity argument for the Test of English as a Foreign Language*. Routledge.
- Chapelle, C., & Voss, E. (Eds.). (2021). *Validity argument in language testing: Case studies of validation research*. Cambridge University Press.
- 中央教育審議会 (2025). 豊かな学びに繋がる学習評価の在り方～過度な負担を生じさせない在り方との両立～. [https://www.mext.go.jp/content/20250711-mxt\\_kyoiku01-000043568\\_03.pdf](https://www.mext.go.jp/content/20250711-mxt_kyoiku01-000043568_03.pdf)
- Dai, D. W., Vu, T., Knoch, U., Lim, A. S., Malone, D. T., & Mak, V. (2024). Expanding Kane's argument-based validity framework: What can validation practices in language assessment offer health professions education? *Medical Education*, 58(12), 1462-1468. <https://doi.org/10.1111/medu.15452>
- 大学入試センター. (2025a). 令和7年度大学入学共通テスト本試験の問題英語(リーディング). [https://www.dnc.ac.jp/albums/abm.php?d=771&f=abm00005680.pdf&n=2025\\_op\\_20\\_reading.pdf](https://www.dnc.ac.jp/albums/abm.php?d=771&f=abm00005680.pdf&n=2025_op_20_reading.pdf)
- 大学入試センター. (2025b). 令和7年度大学入学共通テスト本試験の問題英語(リスニング). [https://www.dnc.ac.jp/albums/abm.php?d=771&f=abm00005676.pdf&n=2025\\_op\\_25\\_listening.pdf](https://www.dnc.ac.jp/albums/abm.php?d=771&f=abm00005676.pdf&n=2025_op_25_listening.pdf)
- Dursun, A., & Li, Z. (2021). A systematic review of argument-based validation studies in the field of language testing (2000–2018). In C. A. Chapelle & E. Voss (Eds.), *Validity argument in language testing: Case studies of validation research* (pp. 45–70). Cambridge University Press.
- 廣森友人・小金丸倫隆 (編). (2024). エンゲージメント×英語授業「やる気」と「意欲」を引き出す授業のつくり方. 明治図書.
- 今井裕之 (2023). 外国語科における三観点評価の課題. *KELES ジャーナル*, 8, 82-86. [https://doi.org/10.18989/keles.8.0\\_82](https://doi.org/10.18989/keles.8.0_82)
- Jin, Y. (2021). Test specifications. In G. Fulcher & L. Harding (Eds.), *The Routledge handbook of language testing* (pp. 271-288). Routledge. <https://doi.org/10.4324/97811003220756-22>
- Jun, H. S. (2014). *A validity argument for the use of scores from a web-search-permitted and web-source-based integrated writing test*. Unpublished doctoral dissertation, Iowa State University.
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112, 527-535. <https://doi.org/10.1037/0033-2909.112.3.527>
- Koizumi, R., Sakai, H., Ido, T., Ota, H., Hayama, M., Sato, M., & Nemoto, A. (2011). Toward validity argument for test interpretation and use based on scores of a diagnostic grammar test for Japanese learners of English. *Japanese Journal for Research on Testing*, 7, 99-119. [https://doi.org/10.24690/jart.7.1\\_99](https://doi.org/10.24690/jart.7.1_99)
- 小泉利恵 (2018). 英語4技能テストの選び方と使い方—妥当性の観点から—. アルク選書.
- 小泉利恵 (編). (2022). 実例でわかる英語スピーキングテスト作成ガイド. 大修館書店.
- 小泉利恵 (2025). 言語テスト開発理論の現在. 全国英語教育学会50回研究大会記念特別誌, 302-305.
- 小泉利恵・阿川敏恵・高橋閑 (2023). 英語4技能テストのスコアレポートを用いた学習の有用性—大学生の反応から—. 清泉女子大学言語教育研究所言語教育研究, 15, 21-43.
- 小泉利恵・印南洋・深澤真 (編). (2017). 実例でわかる英語テスト作成ガイド. 大修館書店.
- 国立教育政策研究所 (2020a). 「指導と評価の一体化」のための学習評価に関する参考資料(小学校編). [https://www.nier.go.jp/kaihatsu/pdf/hyouka/r020326\\_pri\\_gaikokg.pdf](https://www.nier.go.jp/kaihatsu/pdf/hyouka/r020326_pri_gaikokg.pdf)
- 国立教育政策研究所 (2020b). 「指導と評価の一体化」のための学習評価に関する参考資料(中学校編). [https://www.nier.go.jp/kaihatsu/pdf/hyouka/r020326\\_mid\\_gaikokg.pdf](https://www.nier.go.jp/kaihatsu/pdf/hyouka/r020326_mid_gaikokg.pdf)
- 国立教育政策研究所 (2021). 「指導と評価の一体化」のための学習評価に関する参考資料(高等学校編). [https://www.nier.go.jp/kaihatsu/pdf/hyouka/r030820\\_hig\\_gaikokugo.pdf](https://www.nier.go.jp/kaihatsu/pdf/hyouka/r030820_hig_gaikokugo.pdf)
- 駒井健吾 (2024). 観点別評価は言語活動の評価デザインとして適切か. 全国英語教育学会 第49回福岡大会. 2024年8月25日.
- 久保佑輔 (2022). 妥当性検証の枠組みに基づく批判的思考テストの基準関連妥当性の調査—批判的思考力と態度の関係性—. 外国語教育メディア学会中部支部外国語教育基礎研究部会 2022年度報告論集.
- Kumazawa, T., Shizuka, T., & Mochizuki, M. (2016). Validity argument for the VELC Test® score interpretations and uses. *Language Testing in Asia*, 6(2). <https://doi.org/10.1186/s40468-015-0023-3>
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 243-56. <https://doi.org/10.1177/026553229601300302>
- 文部科学省 (2017a). 小学校学習指導要領(平成29年告示)解説(外国語活動・外国語編). [https://www.mext.go.jp/content/20220614-mxt\\_kyoiku02-10002607\\_11.pdf](https://www.mext.go.jp/content/20220614-mxt_kyoiku02-10002607_11.pdf)
- 文部科学省 (2017b). 中学校学習指導要領(平成29年告示)解説(外国語編). [https://www.mext.go.jp/content/20210531-mxt\\_kyoiku01-100002608\\_010.pdf](https://www.mext.go.jp/content/20210531-mxt_kyoiku01-100002608_010.pdf)

## 引用文献

- 文部科学省 (2018). 高等学校学習指導要領(平成30年告示)解説(外国語編). [https://www.mext.go.jp/content/1407073\\_09\\_1\\_2.pdf](https://www.mext.go.jp/content/1407073_09_1_2.pdf)
- 村越亮治・江原大明 (2015). 高校現場における「英語テスト」と「観点別評価」の課題. 神奈川県立国際言語文化アカデミア紀要, 4, 1-18. [https://doi.org/10.20686/academiakiyou.4.0\\_1](https://doi.org/10.20686/academiakiyou.4.0_1)
- Park, M. (2021). Domain Definition Inference for a Virtual Interactive Aviation English Test (VIAET) for Military Air Traffic Controllers. In C. A. Chapelle & E. Voss (Eds.), *Validity Argument in Language Testing: Case Studies of Validation Research* (pp. 73-95). Cambridge University Press.
- 渡邊大志・大鋸雄介・佐藤扶・鶴亀恵一・斉田智里 (2025). 神奈川県公立高等学校入学試験英語問題の分析—観点別学習状況評価の観点から—. *教育デザイン研究*, 16, 146-155. <https://ynu.repo.nii.ac.jp/record/2001651/files/16-1-17.pdf>
- 山本大貴・近藤暁子・鳴海智之・多田ウェンディ・吉田達弘 (2021). 日本の高等学校での活用を想定したペア型スピーキングテストの開発. *外国語教育メディア学会中部支部紀要*, 58, 99-128. [https://doi.org/10.24539/let.58.0\\_99](https://doi.org/10.24539/let.58.0_99)

## 資料1: 教師質問紙調査の指示文

## 高等学校英語科における定期試験・パフォーマンステストの評価に関するアンケート

B I U ☞ ✎

本研究は、高等学校英語科における定期試験およびパフォーマンステストに対する教師の認識を調査し、より妥当性のある学習評価を実施するための改善策を検証することを目的としています。この研究で得られた情報は研究者のみが所持し、あなたの個人情報が漏洩されることはありません。また、あなたの個人情報や個人のデータ等の開示を希望される際はそれに応じます。

以下に自身のメールアドレスを入力する欄があります。メールアドレスの回答は自身の回答の確認用ですので、「同意する」と回答していただいた方でもメールアドレスの情報は使用いたしません。

ご協力のほど、よろしくお願いいたします。

## 資料2: 教師質問紙調査の質問項目と調査結果

|    |                                                                          | 指導経験(年) |      |     | M   | SD  |
|----|--------------------------------------------------------------------------|---------|------|-----|-----|-----|
|    |                                                                          | 0~5     | 6~10 | 11~ |     |     |
| Q1 | 高等学校での英語指導歴                                                              | —       | —    | —   | —   | —   |
| Q2 | 令和6年度にご担当された学年(該当するものを選んでください:複数回答可)                                     | —       | —    | —   | —   | —   |
| Q3 | 「定期試験で測る英語力」について、学校や学年として、<br>「単元の指導と評価の計画」等に明示されていると思いますか?              | 3.0     | 4.5  | 5.3 | 4.3 | 1.6 |
| Q4 | 「パフォーマンステストで測る英語力」について、学校や学年として、<br>「単元の指導と評価の計画」等に明示されていると思いますか?        | 3.0     | 4.0  | 5.3 | 4.1 | 1.5 |
| Q5 | あなたは観点別評価における「読むこと」の「知識・技能」・「思考・判断・表現」の<br>評価規準の違いについて、どの程度理解していると感じますか? | 4.0     | 4.0  | 3.3 | 3.8 | 0.9 |
| Q6 | あなたは観点別評価における「聞くこと」の「知識・技能」・「思考・判断・表現」の<br>評価規準の違いについて、どの程度理解していると感じますか? | 3.3     | 3.3  | 3.3 | 3.3 | 0.8 |
| Q7 | あなたは観点別評価における「書くこと」の「知識・技能」・「思考・判断・表現」の<br>評価規準の違いについて、どの程度理解していると感じますか? | 4.0     | 3.5  | 3.7 | 3.7 | 1.0 |
| Q8 | あなたは観点別評価における「話すこと」の「知識・技能」・「思考・判断・表現」の<br>評価規準の違いについて、どの程度理解していると感じますか? | 2.7     | 3.0  | 3.7 | 3.1 | 1.1 |

## (資料2 続き)

|     |                                                                                                | 指導経験(年) |      |     | M   | SD  |
|-----|------------------------------------------------------------------------------------------------|---------|------|-----|-----|-----|
|     |                                                                                                | 0～5     | 6～10 | 11～ |     |     |
| Q9  | ご自身の勤務校ではCan-doリストを設定していますか?                                                                   | —       | —    | —   | —   | —   |
| Q10 | 令和6年度にご自身をご担当された科目の「定期試験」は、学習指導要領の目標、Can-doリスト、教科書の内容、単元の指導と評価の計画等に基づいて適切に設計されていると思いますか?       | 3.3     | 3.8  | 4.0 | 3.7 | 1.5 |
| Q11 | 令和6年度にご自身をご担当された科目の「パフォーマンステスト」は、学習指導要領の目標、Can-doリスト、教科書の内容、単元の指導と評価の計画等に基づいて適切に設計されていると思いますか? | 3.0     | 4.3  | 4.3 | 3.9 | 1.5 |
| Q12 | 令和6年度にご自身をご担当された科目の「定期試験」が、学習指導要領の目標を適切に反映していると思いますか?                                          | 3.7     | 3.5  | 4.0 | 3.7 | 1.1 |
| Q13 | (Can-doリストを設定していると回答した場合) 令和6年度にご自身をご担当された科目の「定期試験」が、Can-doリストを適切に反映していると思いますか?                | 2.5     | 3.5  | 5.0 | 3.5 | 1.5 |
| Q14 | 令和6年度にご自身をご担当された科目の「定期試験」が、単元の指導と評価の計画を適切に反映していると思いますか?                                        | 3.3     | 3.5  | 4.0 | 3.6 | 1.0 |
| Q15 | 令和6年度にご自身をご担当された科目の「定期試験」が、教科書の内容を適切に反映していると思いますか?                                             | 4.7     | 4.5  | 4.7 | 4.6 | 0.7 |
| Q16 | 令和6年度にご自身をご担当された科目の「パフォーマンステスト」が、学習指導要領の目標を適切に反映していると思いますか?                                    | 4.3     | 4.0  | 4.7 | 4.3 | 1.0 |
| Q17 | (Can-doリストを設定していると回答した場合) 令和6年度にご自身をご担当された科目の「パフォーマンステスト」が、Can-doリストを適切に反映していると思いますか?          | 3.5     | 3.8  | 5.0 | 3.7 | 1.7 |
| Q18 | 令和6年度にご自身をご担当された科目の「パフォーマンステスト」が、単元の指導と評価の計画を適切に反映していると思いますか?                                  | 3.3     | 3.8  | 4.7 | 3.9 | 1.2 |
| Q19 | 令和6年度にご自身をご担当された科目の「パフォーマンステスト」が、教科書の内容を適切に反映していると思いますか?                                       | 4.7     | 4.5  | 5.0 | 4.7 | 1.2 |
| Q20 | 定期試験の設計プロセスにおいて、どのような資料を参照していますか?                                                              | —       | —    | —   | —   | —   |
| Q21 | パフォーマンステストの設計プロセスにおいて、どのような資料を参照していますか?                                                        | —       | —    | —   | —   | —   |
| Q22 | 定期試験の問題は、授業で扱う内容を反映していると思いますか?                                                                 | 5.0     | 5.0  | 5.0 | 5.0 | 1.1 |
| Q23 | パフォーマンステストの内容は、授業で行う活動と一致していると思いますか?                                                           | 5.3     | 4.5  | 4.3 | 5.0 | 1.2 |
| Q24 | 定期試験を作成・実施する際に難しいと感じる点について、自由にご意見をお聞かせください。                                                    | —       | —    | —   | —   | —   |
| Q25 | パフォーマンステストを作成する際に難しいと感じる点について、自由にご意見をお聞かせください。                                                 | —       | —    | —   | —   | —   |
| Q26 | 定期試験やパフォーマンステストを作成・実施する際に工夫している点(例: 生徒の学習につながるように意図している点)について、ご記入ください。                         | —       | —    | —   | —   | —   |
| Q27 | 評価を実施する際に、「この点は適切にできているのか?」と疑問を感じる点があればご記入ください。                                                | —       | —    | —   | —   | —   |

## 資料3: 送付した単元の指導と評価の計画(一部)

| 教科             | 外国語科                                                                                                    | 科目                                                                             | 英語コミュニケーションⅢ                                        | 単位数                                         | 4              |
|----------------|---------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|-----------------------------------------------------|---------------------------------------------|----------------|
| Lesson 1       |                                                                                                         |                                                                                |                                                     |                                             |                |
|                | 内容のまとまりごとの評価規準                                                                                          |                                                                                |                                                     |                                             | 主な<br>評価<br>方法 |
|                | 知識・技能                                                                                                   | 思考・判断・表現                                                                       | 主体的に<br>学習に取り組む態度                                   |                                             |                |
| ・読むこと<br>聞くこと  | [知識] 前置詞を伴う関係代名詞、動名詞の意味上の主語、受け身の完了形、倒置、付帯状況の with を理解している。                                              | 錯視が日常生活に使われている方法について理解できる。                                                     | 錯視が日常生活に使われている方法について自ら理解しようとしている。                   | ペーパー<br>テスト<br><br>振り返りの<br>記述<br>活動の観<br>察 |                |
|                | [技能] 教科書本文の音声を読んだり、本文を読んだりして、本文中の前置詞を伴う関係代名詞、動名詞の意味上の主語、受け身の完了形、倒置、付帯状況の with が使われた英文の意味を把握する技能を身につけている |                                                                                |                                                     |                                             |                |
| (やり取り)<br>話すこと | [知識]                                                                                                    | 理解した内容を自分の言葉で表現し、相手に伝えることができる。また、錯視が日常生活に使われている方法について自分なりに意見を持ち、それを表現することができる。 | 本文から学んだ錯視が日常生活に使われている方法について自分の意見を持ち、積極的に発信しようとしている。 |                                             |                |
|                | [技能]                                                                                                    |                                                                                |                                                     |                                             |                |
| (発表)<br>話すこと   | [知識]                                                                                                    | 理解した内容を自分の言葉で表現し、相手に伝えることができる。また、錯視が日常生活に使われている方法について自分なりに意見を持ち、それを表現することができる。 | 本文から学んだ錯視が日常生活に使われている方法について自分の意見を持ち、積極的に発信しようとしている。 | 個人発表<br><br>振り返りの<br>記述<br>活動の観<br>察        |                |
|                | [技能] 前置詞を伴う関係代名詞、動名詞の意味上の主語、受け身の完了形、倒置、付帯状況の with を用いて一つのテーマについて書き、それを発表する技能を身につけている。                   |                                                                                |                                                     |                                             |                |
| 書くこと           | [知識]                                                                                                    | 理解した内容を自分の言葉で表現し、相手に伝えることができる。また、錯視が日常生活に使われている方法について自分なりに意見を持ち、それを表現することができる。 | 本文から学んだ錯視が日常生活に使われている方法について自分の意見を持ち、積極的に発信しようとしている。 | ワークシ<br>ート<br><br>振り返りの<br>記述<br>活動の観<br>察  |                |
|                | [技能] 前置詞を伴う関係代名詞、動名詞の意味上の主語、受け身の完了形、倒置、付帯状況の with を用いて一つのテーマについて書き、それを発表する技能を身につけている。                   |                                                                                |                                                     |                                             |                |

記録に残す評価 (○)

| 時間          | 主な学習活動                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 知 | 思 | 主 |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|---|---|
| 1<br>5<br>6 | <ul style="list-style-type: none"> <li>単元の目標や進め方を確認し、学習の見通しをもつ。</li> <li>提示された質問を基に自分について考える。</li> <li>錯視が日常生活に使われている方法について聞いたり、読んだりし、理解する。</li> <li>教科書の内容や自分の考えを自分の言葉で他に伝えたり、書いたりする。</li> <li>本文で使われている前置詞を伴う関係代名詞、動名詞の意味上の主語、受け身の完了形、倒置、付帯状況の with の使い方を理解し、身につける。</li> <li>教科書の内容を応用し、自分の考えを積極的に自分の言葉で他に伝えたり、書いたりしようとする。</li> <li>本文で使われている前置詞を伴う関係代名詞、動名詞の意味上の主語、受け身の完了形、倒置、付帯状況の with の使い方を理解し、積極的に応用場面で使おうとする。</li> <li>上記の振り返りを Booklet に記入する。</li> </ul> | ○ | ○ | ○ |
| 後<br>日      | <ul style="list-style-type: none"> <li>ペーパーテスト<br/>初めて聞いたり読んだりする英文からも出題されます。</li> </ul>                                                                                                                                                                                                                                                                                                                                                                           | ○ | ○ |   |

活動状況を見取り、学習や指導に生かすための評価は毎時間行います。

1. 3年間の計画（令和4年設定）

目標：多様な他者と協働して新たな価値を創出し、社会に貢献する科学的探究者の育成

4 観点（5 領域）の目標

|                            | 目標                                                                                                  | 1 年                                                                                                                                                  | 2 年                                                                                                                              | 3 年                                                                                                                                    |
|----------------------------|-----------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| READING                    | 英語で書かれた様々な文章に触れ、多様な考え方・ものの見方を理解し、自分の考えを持つことができる。<br>(教科書でも多様なジャンルを扱う。共創探究で論文にも触れさせたい)               | ジャンルの意識しリーディングスキルを使い分けることができる。パラグラフ構成について理解して筆者の主張を掴むことができる。WPM は 100 を目指す。                                                                          | ディスコースマーカーや結束性(cohesion)、一貫性(coherence)などを意識して、パラグラフのつながりや文章全体を理解できる。WPM は 130 を目指す。(共通テストで余裕があるレベル)                             | 論文なども含め、Authentic な英文を読むことができる。WPM は 150 を目指す。(TOEIC ですべて読み終わるレベル)                                                                     |
| WRITING                    | コミュニケーションの相手やジャンルに応じた適切な表現を用いて、自分の考えを伝えることができる。<br>(共創でアブストラクト→パラグラフライティング、英検・外語大・IELTS→エッセイライティング) | 英文を書くことに抵抗を感じない。<br>(Fluency, Global/Local Error)<br>基本的な型(理由の述べ方、パラグラフライティング)を理解しそれに従って英文を書くことができる。ジャンルや読み手を意識してライティングをすることができ 100 語程度の文量を書くことができる。 | 英文の細かい部分まで意識できる。<br>(Accuracy, Local Error にも注意)<br>内容の深化を図る。(論理的に考えて文章が構成されているか)<br>200 語程度書くことができる。                            | genre (手紙やメールなど) や receiver (友達や上司など)に応じて適切な表現や幅広い表現 (無生物主語、分詞構文など)を用いて英文を書くことができる。<br>20 分で 200 語程度 (分量最大 IELTS の 250 語)の英文を書くことができる。 |
| LISTENING                  | 様々な英語に触れ、要点などを押さえることができる。<br>(海外の研究発表を聞けるように)                                                       | 身近なテーマ (自己紹介など) や最新の話題について聞き取り、理解ができる。<br>話される内容をメモしながら聞くことができる。<br>実際の英語に慣れることができる。                                                                 | 社会的なテーマ (学術的なスピーチなど) や最新の話題について聞き取り、理解ができる。1 分程度の実際の英語を聞き取ることができる。                                                               | 社会的なテーマ (英語での専門的な授業など) や最新の話題について聞き取り、理解ができる。また、複数の話者や雑音が入っているものも聞き取ることができる。                                                           |
| SPEAKING<br>(INTERACTION)  | 英語を話すことに対して抵抗を感じず、積極的に他者とコミュニケーションを図ることができる。聞き手に応じた適切な表現を用いて、自分の考えを伝えたり、質問したりできる。<br>(海外の研究発表の質疑応答) | 英語を話すことに抵抗を感じない。(Fluency) Filler や適切な表現を使いながら会話をつなげることができる。<br>自分の周りのことについて 1 分間話すことができる。(CEFR A2 レベル) 疑問文を作ることができる。                                 | 発話の細かい部分まで意識できる、Error を減らしていく。(Accuracy)<br>聞き手からの質問など(+1)があっても自然な回答や会話を続けることができる。<br>自分の周りの社会のことについて 1 分間話すことができる。(CEFR B1 レベル) | register や genre, context を意識して会話を続けることができる。(Accuracy)<br>自然な流れで会話の目的を達成することができる。<br>様々な内容について 1 分間話すことができる。(CEFR B2 レベル)              |
| SPEAKING<br>(PRESENTATION) | 聞き手に応じた適切な表現やアイコンタクトなどの工夫を生かして、自分の考えを伝えることができる。<br>(ノンバーバルとバーバルを育成する)                               | 聞き手を意識して発表をすることができる。ノンバーバルの視点の定着。                                                                                                                    | 聞き手が理解しやすいため発表をすることができる。定型句などを用いての発表ができる。バーバルの視点の定着。                                                                             | 聞き手を意識して理解しやすいため発表をすることができる。グラフの説明やデータを用いた研究を英語で行うことができる。                                                                              |
| 意識している内容・活動                | ・アカデミックス点での資質・能力の育成<br>→海外での発表に耐えうる資質・能力の育成<br>例) 論文の読解、発表や質問を聞く<br>論文を書く、発表や質問に対する応答               | R: 多様なジャンルの読解<br>W: パラグラフ、ジャンルの読解<br>L: メモの取り方、音の特徴 (脱落など) に慣れる<br>S: 聞き手の意識、ストラテジー、ノンバーバル                                                           | R: まとまった分量・論文 (一部) の読解<br>W: エッセイライティング<br>L: 様々な状況のリスニングに慣れる<br>S: 聞き手の意識、ストラテジー、バーバル                                           | R: 多様なジャンルを大量に読む、論文の読み取り<br>W: エッセイライティング<br>L: まとまった分量の聞き取り<br>S: グラフなども含めてバーバル、質問を作る                                                 |

## 資料5: 生徒質問紙調査の指示文

## 高等学校英語科における定期試験に関するアンケート

**B** *I* U ☞ ✕

このアンケートは、みなさんが英語科の定期試験についてどのように感じているかを調査するものです。みなさんの声を集めて、これからの英語の学習やテストが、さらに良くなるように役立てたいと考えています。

このアンケートで集めた情報は、研究目的でのみ使用します。公表に際しては情報提供者のプライバシーが侵害されないように最大限の注意を払い、個人名が特定されないように配慮いたします。

また、このアンケートに答えたことで成績などに影響が出ることもありません。安心して、自分の考えを書いてください。

アンケートの内容について、知りたいことや確認したいことがあれば、英語科：大鋸（おおが）まで声をかけてください。

ご協力のほど、よろしくお願いいたします。

資料6: 生徒質問紙調査の質問項目及び調査結果 .....

|     |                                                  | 1年  |     | 2年  |     | 3年  |     | 全体  |     |
|-----|--------------------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
|     |                                                  | M   | SD  | M   | SD  | M   | SD  | M   | SD  |
| Q1  | 英語の学習を始めたのはいつですか?                                | —   | —   | —   | —   | —   | —   | —   | —   |
| Q2  | 英語圏の国に留学または滞在したことがありますか?                         | —   | —   | —   | —   | —   | —   | —   | —   |
| Q3  | (「はい」の場合)どのくらいの期間滞在しましたか?                        | —   | —   | —   | —   | —   | —   | —   | —   |
| Q4  | 英語コミュニケーションの授業では、主にどのように活動していますか?(複数回答可)         | —   | —   | —   | —   | —   | —   | —   | —   |
| Q5  | 英語コミュニケーションの授業では、先生が話し、生徒が聞く時間が多い。               | 3.1 | 1.1 | 3.8 | 1.3 | 3.8 | 1.2 | 3.5 | 1.2 |
| Q6  | 英語コミュニケーションの授業では、先生は主に英語を話す。                     | 3.9 | 1.2 | 3.9 | 1.1 | 4.1 | 1.1 | 3.9 | 1.2 |
| Q7  | 英語コミュニケーションの授業では、先生と話すとき、生徒はほとんど英語を使う。           | 2.9 | 1.1 | 3.0 | 1.2 | 2.7 | 1.1 | 2.9 | 1.1 |
| Q8  | 英語コミュニケーションの授業中、生徒同士で話すとき、ほとんど英語を使う。             | 3.4 | 1.2 | 3.3 | 1.2 | 2.8 | 1.1 | 3.2 | 1.2 |
| Q9  | 英語コミュニケーションの授業のスピーキング活動では、即興で英語を話すことが求められることが多い。 | 3.9 | 1.2 | 3.6 | 1.2 | 3.6 | 1.2 | 3.7 | 1.2 |
| Q10 | 英語コミュニケーションの授業の活動は主に何に焦点を当てていますか?(複数回答可)         | —   | —   | —   | —   | —   | —   | —   | —   |
| Q11 | 英語コミュニケーションの授業では何に多くの時間を使いますか?(複数回答可)            | —   | —   | —   | —   | —   | —   | —   | —   |
| Q12 | 英語コミュニケーションの授業では定期試験対策の学習を多く行っている。               | 3.7 | 1.2 | 3.3 | 1.1 | 3.3 | 1.3 | 3.5 | 1.2 |
| Q13 | 1週間にどのくらいの時間、英語の宿題をしますか?                         | —   | —   | —   | —   | —   | —   | —   | —   |
| Q14 | 学校の宿題の内容は主に何ですか?(複数回答可)                          | —   | —   | —   | —   | —   | —   | —   | —   |
| Q15 | 高校の英語コミュニケーションの授業に満足している。                        | 4.5 | 1.1 | 3.8 | 1.3 | 3.8 | 1.3 | 4.1 | 1.3 |
| Q16 | 高校の英語コミュニケーションの授業によって、どのスキルが向上したと思いますか?(複数回答可)   | —   | —   | —   | —   | —   | —   | —   | —   |
| Q17 | 高校の英語コミュニケーションの授業について、他にコメントがありますか?              | —   | —   | —   | —   | —   | —   | —   | —   |
| Q18 | 今回の定期試験は楽しかった。                                   | 2.8 | 1.5 | 2.6 | 1.3 | 2.8 | 1.4 | 2.8 | 1.4 |
| Q19 | 今回の定期試験は難しかった。                                   | 3.9 | 1.3 | 4.3 | 1.1 | 4.1 | 1.1 | 4.1 | 1.2 |
| Q20 | 今回の定期試験の指示は明確だった。                                | 4.5 | 1.2 | 4.5 | 1.1 | 4.3 | 1.3 | 4.4 | 1.2 |
| Q21 | 今回の定期試験の試験時間は十分だった。                              | 4.0 | 1.5 | 3.5 | 1.4 | 3.1 | 1.4 | 3.6 | 1.5 |
| Q22 | 今回の試験時間のために学校外でのテスト準備に何時間費やしましたか?                | —   | —   | —   | —   | —   | —   | —   | —   |
| Q23 | 今回の定期試験の詳細について知ったとき、どのように試験の準備をしましたか?            | —   | —   | —   | —   | —   | —   | —   | —   |
| Q24 | 今回の定期試験のために学校外では主に何を中心に学習しましたか?(複数回答可)           | —   | —   | —   | —   | —   | —   | —   | —   |
| Q25 | 学校の授業は今回の定期試験の準備に役立った。                           | 4.4 | 1.1 | 4.0 | 1.1 | 3.8 | 1.3 | 4.1 | 1.2 |
| Q26 | 今回の定期試験の問題は学校の授業で行った内容と似ていた。                     | 4.6 | 1.1 | 4.2 | 1.1 | 3.9 | 1.2 | 4.3 | 1.2 |
| Q27 | 今回のテストは、定期試験として適している。                            | 4.8 | 1.0 | 4.4 | 1.1 | 4.4 | 1.1 | 4.6 | 1.1 |
| Q28 | 今回の定期試験後、自信がついたものはどれですか?(複数回答可)                  | —   | —   | —   | —   | —   | —   | —   | —   |
| Q29 | 今回の定期試験後、さらに練習したいと思ったものはどれですか?(複数回答可)            | —   | —   | —   | —   | —   | —   | —   | —   |
| Q30 | スコアレポート(得点個表)は自身の強みや弱点を理解するのに役立つと思う。             | 5.0 | 1.0 | 4.4 | 1.2 | 4.4 | 1.3 | 4.6 | 1.2 |
| Q31 | スコアレポート(得点個表)は明確である。                             | 5.0 | 1.0 | 4.5 | 1.0 | 4.4 | 1.2 | 4.6 | 1.1 |
| Q32 | スコアレポート(得点個表)は解釈しやすい。                            | 4.9 | 1.1 | 4.5 | 1.0 | 4.3 | 1.2 | 4.6 | 1.1 |
| Q33 | 授業で受けた指導は、今回の定期試験を受けるのに十分だったと思う。                 | 4.7 | 1.1 | 4.2 | 1.1 | 4.0 | 1.2 | 4.4 | 1.2 |
| Q34 | そう思う理由を書いてください。                                  | —   | —   | —   | —   | —   | —   | —   | —   |

## 資料7: 生徒質問紙調査の選択式回答項目の選択肢及び回答結果

|     |                                              | 回答人数<br>(%)            |
|-----|----------------------------------------------|------------------------|
| Q1  | 英語の学習を始めたのはいつですか?                            | 0～6歳                   |
|     |                                              | 7～12歳                  |
|     |                                              | 13～16歳                 |
| Q2  | 英語圏の国に留学または滞在したことがありますか?                     | はい                     |
|     |                                              | いいえ                    |
| Q3  | ([「はい」の場合]どのくらいの期間滞在しましたか?)                  | 1ヶ月未満                  |
|     |                                              | 1～12ヶ月                 |
|     |                                              | 12ヶ月以上                 |
| Q4  | 英語コミュニケーションの授業では主にどのように活動していますか?<br>(複数回答可)  | 個人                     |
|     |                                              | グループ                   |
|     |                                              | ペア                     |
|     |                                              | クラス全体                  |
| Q10 | 英語コミュニケーションの授業の活動は主に何に焦点を当てていますか?<br>(複数回答可) | 語彙                     |
|     |                                              | 文法                     |
|     |                                              | 発音                     |
|     |                                              | リスニング                  |
|     |                                              | リーディング                 |
|     |                                              | スピーキング                 |
|     |                                              | ライティング                 |
| Q11 | 英語コミュニケーションの授業では何に多くの時間を使いますか?<br>(複数回答可)    | 語彙                     |
|     |                                              | 文法                     |
|     |                                              | 発音                     |
|     |                                              | テキストを読んで質問に答える         |
|     |                                              | モノローグを聞いて質問に答える        |
|     |                                              | ダイアログを聞いて質問に答える        |
|     |                                              | メールを書く                 |
|     |                                              | 記事を書く                  |
|     |                                              | 物語を書く                  |
|     |                                              | エッセイを書く                |
|     |                                              | レビューを書く                |
|     |                                              | ペアで身近な話題について質問・回答する    |
|     |                                              | 一人で絵を説明する              |
|     |                                              | ペア・グループで絵について話し合う      |
|     |                                              | ペア・グループでアイデアやトピックを議論する |
|     |                                              | ペア・グループで意思決定を行う        |
| Q13 | 1週間にどのくらいの時間、英語の宿題をしますか?                     | 1時間未満                  |
|     |                                              | 1～2時間                  |
|     |                                              | 2～3時間                  |
|     |                                              | 4時間以上                  |

## (資料7 続き)

|     |                                                    |        | 回答人数<br>(%) |
|-----|----------------------------------------------------|--------|-------------|
| Q14 | 学校の宿題の内容は主に何ですか? (複数回答可)                           | 語彙     | 417(45.1%)  |
|     |                                                    | 文法     | 273(29.5%)  |
|     |                                                    | 発音     | 110(11.9%)  |
|     |                                                    | リスニング  | 182(19.7%)  |
|     |                                                    | リーディング | 278(30.1%)  |
|     |                                                    | スピーキング | 239(25.9%)  |
|     |                                                    | ライティング | 226(24.5%)  |
| Q16 | 高校の英語コミュニケーションの授業によって、どのスキルが向上したと思いますか?<br>(複数回答可) | リスニング  | 284(45.6%)  |
|     |                                                    | リーディング | 421(42.9%)  |
|     |                                                    | ライティング | 190(20.6%)  |
|     |                                                    | スピーキング | 396(30.7%)  |
| Q22 | 今回の試験時間のために学校外でのテスト準備に何時間費やしましたか?                  | 0時間    | 135(14.6%)  |
|     |                                                    | 1～5時間  | 636(68.8%)  |
|     |                                                    | 6～10時間 | 120(13.0%)  |
|     |                                                    | 11時間以上 | 33(3.6%)    |
| Q24 | 今回の定期試験のために学校外では主に何を中心に学習しましたか?<br>(複数回答可)         | 語彙     | 549(59.4%)  |
|     |                                                    | 文法     | 384(41.6%)  |
|     |                                                    | 発音     | 40(4.3%)    |
|     |                                                    | リスニング  | 81(8.8%)    |
|     |                                                    | リーディング | 459(49.7%)  |
|     |                                                    | スピーキング | 56(6.1%)    |
|     |                                                    | ライティング | 102(11.0%)  |
|     |                                                    | ない     | 74(8.0%)    |
| Q28 | 今回の定期試験後、自信がついたものはどれですか?<br>(複数回答可)                | 語彙     | 231(25.0%)  |
|     |                                                    | 文法     | 224(24.2%)  |
|     |                                                    | 発音     | 32(3.5%)    |
|     |                                                    | リスニング  | 71(7.7%)    |
|     |                                                    | リーディング | 269(29.1%)  |
|     |                                                    | スピーキング | 27(2.9%)    |
|     |                                                    | ライティング | 37(4.0%)    |
|     |                                                    | ない     | 291(31.5%)  |
| Q29 | 今回の定期試験後、さらに練習したいと思ったものはどれですか?<br>(複数回答可)          | 語彙     | 533(57.7%)  |
|     |                                                    | 文法     | 417(45.1%)  |
|     |                                                    | 発音     | 108(11.7%)  |
|     |                                                    | リスニング  | 244(26.4%)  |
|     |                                                    | リーディング | 536(58.0%)  |
|     |                                                    | スピーキング | 128(13.9%)  |
|     |                                                    | ライティング | 167(18.1%)  |
|     |                                                    | ない     | 39(4.2%)    |

## 第37回 研究助成

## A 研究部門・報告Ⅲ・英語能力テストに関する研究

# 生成 AI を用いたループリック評価の 妥当性と可能性の探究: 教員評価と生成 AI による評価の比較研究

研究者: 大和田 彩 高知県/高知県立室戸高等学校 教諭(申請時: 米国/ハワイ大学マノア校大学院 在籍)

《研究助言者: 竹内 理》

## 概要

本研究の目的は、高校生の英作文に対する従来の教員による評価(以下、教員評価)と、生成 AI (ChatGPT) による評価結果を比較し、生成 AI を活用したループリック評価の妥当性と可能性を検証することである。近年、生成 AI を教育評価に活用することで、評価の効率化や公平性の向上が期待されているが、その評価結果が教員評価とどの程度一致するのか、また AI 評価に特有の傾向が存在するのかについては、十分に検証されていない。本研究では、高校生30名の英作文を対象に、教員20名による評価と ChatGPT の評価を比較した。分析の結果、文法・語彙・構成といった形式的観点では AI と教員の評価が高い一致度を示すことが確認された一方で、主体的態度といった主観的観点では、AI は教員よりも過小評価する傾向がみられた。さらに、Technology Acceptance Model (TAM) に基づくアンケートでは、教員は AI の有用性や効率性を一定程度認めつつも、「主観的観点の判断は人間が担うべき」とする慎重な姿勢が示された。

## 1 はじめに

英作文は、英語教育において学習者の論理的思考力や表現力、さらには批判的思考力を育成するための重要な学習活動のひとつである。単なる語彙や文法の習得にとどまらず、自分の考えを筋道立てて表現し、相手に伝えるという実践的な力を育てる場として、英作文は中等教育から高等教育にかけて幅広く導入されている。その一方で、評価の在り方については長年にわたり課題が指摘されてきた。Bloom et al. (1956) の教育目標分類で示されているように、英作文の評価には知識や理解だけでなく、応用、分析、統合といった高次の認知的プロセスが関与しており、単純な正誤判定では捉えきれない複雑な判断が求められる。従来の英作文評価は、教員の専門的判断に基づいて行われることが多く、その専門性は評価の質を担保する一方で、評価者間の一貫性や公平性を確保するうえでの大きな障壁ともなってきた。特に、大規模なクラス編成や多忙な学校現場においては、すべての生徒に対して十分な時間を割いた丁寧な評価とフィードバックを行うことは容易ではない。Ellis et al. (2006) は、こうした教育実践において、リキャストのような暗示的フィードバックよりも、メタ言語的説明を伴う明示的フィードバックの方が学習者の理解を深めやすいと指摘している。中でも、ループリックやコメントを活用した訂正フィードバック (Written Corrective Feedback: WCF) は、誤りの明示と修正を通じて言語習得を促進する有効な手法とされ、実際の授業においても高く評価されている (Zeevy-Solovey, 2024)。しかしながら、すべての教員がこのような包括的なフィードバックを実践するには、人的、時間的リソースの限界があるのが現実である。大人数のクラスで一人ひとりの作文を丁寧に読み、誤りを指摘し、改善の方向性を提示する作業は、多く

の教員にとって大きな負担となっている (Lin & Crosthwaite, 2024; Shermis et al. 2002)。このような状況において、学習者にとっても、必要なタイミングで十分なフィードバックを受け取る機会が限られてしまうという課題が生じている。

こうした背景を受けて、近年では教育現場における評価支援ツールとしてのAI技術の活用が注目されている。とりわけ、言語学習支援を目的とした生成AIの進展は目覚ましく、自然言語処理能力の向上により、質の高いフィードバックを瞬時に提供することが可能となりつつある (ElEbyary et al., 2024)。生成AIは、学習者の英作文に対して個別かつ即時的な応答を生成し、学習者の理解を深め、主体的な改善行動を促すツールとしての可能性を秘めている。文部科学省(2024)も、生成AIを「伴走者」として位置づけ、教員の負担軽減と学習者の主体性育成の両立を図る補助的な活用を推奨しており、その導入に対する社会的認知も進みつつある。さらに、Technology Acceptance Model(TAM)の理論的枠組みによれば、新しい技術が教育現場に受け入れられるかどうかは、有用性と使いやすさによって大きく左右されるとされている (Davis, 1989; Chocarro et al., 2023)。つまり、どれほど技術的に優れたツールであっても、教員や学習者とその利便性や必要性を実感できなければ、教育実践への定着は困難である。したがって、AI導入に際しては、技術的条件のみならず、教育関係者の心理的な受容性に配慮した支援や研修の整備が不可欠であるといえる。

以上を踏まえ、本研究では、日本の高等学校における英作文指導の実態に即して、高校生による英作文に対する教員評価と、生成AI(ChatGPT)による評価結果を比較し、AIを活用した英作文評価の妥当性と有効性を検討する。また、教員および学習者のAI評価に対する受容意識を明らかにすることで、今後の教育現場におけるAI導入に向けた具体的なかつ実践的な示唆を提示することを目的とする。

## 2 先行研究

### 2.1 L2ライティングにおけるフィードバックの役割と課題

第二言語習得 (Second Language Acquisition: SLA) に関する研究領域では、学習者の言語運用能力や習得プロセスを説明するために、さまざまな理論的枠組みが提案されてきた。その中でも、行動主義の立場に基づく理論では、「すべての行動は刺激と反応による条件づけの結果である (VanPatten & Williams, 2007)」とされ、学習はさまざまなフィードバックの積み重ねによって形成されると理解されている。この観点は、英作文指導におけるフィードバックの意義と深く関係しており、学習者の言語表現力を高めるための要因として、教員のフィードバックが重要な役割を担っている。フィードバックは、学習者に対して自身のエラーや課題を気づかせ、改善へと導くための具体的な情報を提供するものであり、英作文の完成度を高めるための中心的手段である (Ferris & Roberts, 2001; Troia, 2014)。特に、エラーの種類や学習段階に応じた適切なフィードバックの方法が、学習者の言語習得に与える影響は大きい。Ellis et al. (2006) は、さまざまなフィードバックの形式の中でも、明示的な訂正フィードバックが学習者の注意を引きやすく、エラー修正の定着において高い効果を示すことを明らかにしている。さらに、Teimouri (2017) は、単なる訂正にとどまらず、学習者の努力や長所に光を当てるポジティブフィードバックを併用することが、学習意欲や自信の向上につながると指摘しており、感情的側面への配慮の重要性を示唆している。加えて、Troia (2014) は、フィードバックの在り方について、教員から学習者への一方的なコメントに限定するのではなく、学習者との双方向的なコミュニケーションを通じて意味交渉を行う「対話的フィードバック (dialogic feedback)」の有効性に注目すべきであると述べている。こうしたフィードバックは、単なる誤りの指摘ではなく、学習者が自らの言語表現を見直し、再構成する過程に主体的に関わる機会を提供する点で、従来の指導よりも深い学びを促すことが期待されている。

しかしながら、このように理想的なフィードバックを実現するためには、教員に多くの時間的、精神的

負担がかかるのが現実である。特に、クラスの規模が大きく、一人の教員が複数の英作文に対応しなければならない場合、公平性や一貫性を維持しながら個別にフィードバックを行うことは非常に困難である。また、学習者によって必要とする支援の内容が異なるため、画一的な指導では対応しきれないというジレンマも存在する。このような状況を踏まえると、英作文における効果的なフィードバックを持続的に実施するためには、教員を支える補助的な手段の導入が不可欠であるといえる。近年では、教育技術の進展に伴い、AIをはじめとするテクノロジーを活用した支援ツールが注目を集めており、こうした技術をどのように教育実践に統合し、学習効果の向上につなげていくかが、今後の大きな課題となっている。フィードバックの質と量の両立、学習者の個性への対応、教員の負担軽減といった複合的な課題に対し、テクノロジーの活用がいかなる可能性を持つのかを検証することは、今後の英作文指導の実践にとって極めて重要な視点である。

## 2.2 AIによるライティング評価の可能性と課題

近年、生成AIの急速な技術革新により、言語教育への応用が著しく進展している。従来のAWE(自動作文評価)システムに比べて、ChatGPTのような生成AIは柔軟かつ学習者にとって理解しやすいフィードバックを提示できる点が注目されており(Han & Li, 2024)、特にその即時性と一貫性は、正確かつ効率的なフィードバックの提供に貢献する(Polakova & Ivenz, 2024; Mohammed & Khalid, 2025)。一方で、Teng(2024)の研究によれば、EFL学習者はChatGPTのフィードバックの有効性を認めつつも、人間味や具体性の欠如などの限界を指摘しており、教員による補完の必要性が示唆されている。

AIの評価の活用に関しては、その信頼性と妥当性が実用化に向けた重要な課題とされている。Yang(2024)は、ChatGPTにおける自己内一致信頼性の低さを指摘し、同一基準に基づく安定した評価の困難さを示している。また、Uyar and Büyükahiska(2025)は、AIと人間評価者によるスコアに有意差が見られ、特に創造性や論理構成といった高次スキルにおいてその差が顕著であることを報告している。Yao et al.(2025)は、生成AIを補助ツールとしてL2ライティング指導に組み込む際、教員がどのように活用するか、その実践的統合のプロセスに不確実性があると指摘しており、Yavuz et al.(2025)も、LLM(大規模言語モデル)の出力に過度に依存することで、評価の信頼性が損なわれる可能性を示している。さらに、Steiss et al.(2024)はAIのフィードバック品質は一定水準を維持しているとする一方で、Fuller and Bixby(2024)は、生成AIの使用方法によっては採点やフィードバックにばらつきが生じ、信頼性や一貫性に欠ける可能性があるとは指摘している。

このように、AI評価は即時性や一貫性といった利点を有し、学習者のアウトプットの改善や教員の負担軽減に貢献し得るが、最終的な評価は教員が担うべきであり、AIは補助的手段として活用されるべきである(Li et al., 2024)。その精度を高めるには、明確なルーブリックと評価者間の信頼性(Inter Rater Reliability: IRR)の確保が不可欠である(Kaldaras et al., 2022)。ルーブリックは主にパフォーマンスの評価に活用される(Brookhart, 2013)。また、AIの導入に際しては、教育現場における倫理的配慮や人間中心の価値観を損なわない戦略的運用が求められる(Pozdniakov et al., 2024)。教員によるフィードバックは誤りの特定や内省的思考の促進に有効であり、AIは低次の誤りへの対応に強みを持つ(Fan et al., 2024)。ChatGPTの統合型ライティングタスクへの有効性については今後の検証が必要とされ(Kim et al., 2024)、AIと人間のフィードバック併用による効果もまだ十分に研究されていない(Tran, 2025)。Tate et al.(2024)の分析では、AI採点による明確な悪影響は認められず、プロンプトやコーパスの工夫によって改善の余地があることが示された。したがって、AIと人間による評価の差異および相補的可能性を明らかにすることは、教育現場における効果的かつ持続可能な評価の在り方を再考するうえで、重要な意義を持つ。

## 2.3 AI評価の受容と課題

AIを教育現場に導入するには、単に技術的な精度や機能性の高さのみならず、それを使用する教員や学習者自身の「受容意識」が極めて重要な要素として作用する。つまり、技術が有しているポテンシャルが教育実践において十分に活かされるかどうかは、実際に利用する人々の態度や意識によって大きく左右されるのである。この点について理論的枠組みを提供するのが、Davis(1989)によって提唱された技術受容モデル(Technology Acceptance Model: TAM)である。TAMは、新たな技術が利用者に受け入れられるかどうかを予測するための代表的なモデルであり、「知覚された有用性(Perceived Usefulness: PU)」および「知覚された使いやすさ(Perceived Ease of Use: PEOU)」、「行動意図(Behavioral Intention: BI)」という三つの主要な概念に着目する。ここで言う「有用性」とは、あるシステムやツールを使用することで自身の仕事や学習のパフォーマンスが向上すると利用者が信じる程度を指し、「使いやすさ」とは、そのシステムの操作が困難ではなく、ストレスなく活用できると感じる程度を意味する。これら二つの知覚が、利用者の態度や使用意図(BI)を大きく左右し、最終的な利用行動に結び付くとされている。このモデルは、ICT機器やAIツールといった教育テクノロジーを学校教育へ導入する際にも、極めて有効な理論的枠組みとして多くの研究に採用されている。

さらに、近年の研究ではTAMを発展的に活用し、実際の教育現場におけるAI導入の受容要因についての実証的分析が進められている。例えば、Chocarro et al.(2023)の研究は、教員の年齢や一般的なデジタルスキルよりも、AIツールの「使いやすさ」および「有用性」に対する主観的な認知の方が、AIの活用意図を高めるうえでより強力な予測因子であることを明らかにした。つまり、教員が若いかどうかやICTに熟達しているかよりも、「このツールは役に立ちそうだ」「簡単に使えそうだ」と感じられることが、AI受容の鍵を握っているという指摘である。また、Lai et al.(2023)は、学習者側の動機づけタイプとAIツールの受容との関係に注目し、学習者の内発的、外発的動機づけの違いが、AIの有用性認知や使用意図に異なる形で影響を与えることを示した。具体的には、外発的動機づけ(成績や報酬など外的要因による動機)が強い学習者は、AIの「有用性」に対する評価が使用意図に影響を及ぼしやすく、一方で、内発的動機づけ(知的好奇心や自己成長への欲求)が高い学習者は、「使いやすさ」の認知が使用意図により強く関与することが確認されている。これにより、AIツールを導入する際には、学習者の動機づけの傾向に応じてアプローチを工夫する必要があることが示唆される。

先にも述べたように、文部科学省(2024)は生成AIを教育現場における「伴走者」として位置づけており、教員の業務負担を軽減する一方で、学習者の主体的な学びを支援することが望ましい活用の在り方としている。しかし、このような理想的な活用を実現するためには、教員や学習者がAIによる評価に対して十分な信頼を寄せ、安心して活用できる心理的環境が整っていることが前提条件となる。

## 3 研究, 実践, 調査方法

### 3.1 調査の目的

本研究は、生成AI(ChatGPT)による英作文評価と高等学校英語教員による評価を比較し、その一致度や評価傾向を明らかにするとともに、AI評価の教育的活用の可能性を検討することを目的とする。具体的には、まずAIと教員がどの程度同じ判断を下しているのかを確認する(RQ1)。次に、その一致度を補足的に理解するため、観点別にAIが一貫して過小または過大評価を行う傾向が存在するのかを分析する(RQ2)。これら二つの問いによって、AI評価の特徴を明らかにする。そのうえで教育現場におけるAI評価の導入の可能性を探るために、教員がAI評価をどのように受容し、どの場面で補助的に活用できると考えているのかを検討する(RQ3)。以上を踏まえ、以下の3点を検証課題(Research Questions: RQ)と

して設定した。

RQ1 生成 AI による英作文評価は、教員による評価とどの程度一致するのか。

RQ2 評価観点ごとに、生成 AI に特有の過小または過大評価の傾向は見られるか。

RQ3 RQ1と2で明らかになった一致度や傾向差を前提として、高等学校英語教員は生成 AI による英作文評価をどのように受容しているか。

### 3.2 研究デザインと評価実施手順

本研究は、日本国内の高等学校に勤務する英語教員による英作文評価と、生成 AI (ChatGPT) による観点別評価とを比較することで、それぞれの評価傾向、安定性、実用性について実証的に検討し、AI の補助的活用の可能性と教員の受容要因を明らかにすることを目的として実施された。

評価対象となった英作文は、筆者が勤務する高等学校に在籍する、研究実施当時高校2・3年生の30名が書いたものである。英語力については、CEFR で概ね A2～B1 レベルに相当している。テーマは「あなたが学びたい第三言語とその理由」とし、主張と二つの理由を明確に示す構成で書くように指導を行った。語数の目安は100～120語とし、オンラインを含む辞書の使用は許可したが、翻訳ツールや文法補助機能などの使用は禁止した。提出された英作文はすべて無記名化され、評価用データセットとして整理、保存された。

評価には、英検の採点基準に準じたルーブリックと、新学習指導要領に準拠した観点別ルーブリックの2種類を使用した。英検準拠型では「内容」「構成」「語彙」「文法」の4観点について、それぞれ0～4点の5段階で評価を行い、学習指導要領準拠型では「知識、技能」「思考、判断、表現」「学習に主体的に取り組む態度」の3観点について、a(優れている)～c(改善が必要)の3段階で評価を行った。教員による評価は、全国の高等学校に勤務する英語教員20名を対象として実施した。年齢層は20代から50代まで幅広く、教員経験年数も1年未満から11年以上と多様であった。評価に用いるルーブリックは学習指導要領準拠、学校独自の様式、個人作成、英検基準などさまざまであり、ICT や生成 AI の活用度も積極的な者から未経験の者まで幅広い分布を示した。各教員には30作品のうち6作品を無作為に割り当て、両方のルーブリックに基づいて観点別に評価してもらった。事前に評価マニュアルを配布し、評価者間の一貫性を担保するよう努めた。すべての評価は Google フォームを用いてスコア形式で回収し、自由記述によるフィードバックや文法などの誤りについての修正提案などは求めなかった。一方、AI による評価には、有料版の ChatGPT (GPT-4o) を使用した。それぞれのルーブリックに対応する2種類のカスタム GPT (英検準拠型および学習指導要領準拠型) を作成し、同一の30作品に対して観点別のスコアを出力させた。各英作文については複数回 (最低3回、最高5回) 評価を生成し、出力の整合性を確認したうえで、内容が一致した場合にその結果を最終スコアとして採用した。AI による評価結果はすべてチャット履歴の URL とともに記録、保存され、検証可能な形で管理された。収集された評価結果のうち、複数のスコアが併記されていたものなど一意に解釈できないデータは「分析不能データ」として除外し、最終的に教員評価113件、AI 評価120件を分析対象とした。

### 3.3 評価結果の分析

本研究では、AI と教員による評価結果の傾向や差異、そして評価の一貫性や安定性の比較を通じて、生成 AI による英作文評価の有効性と限界を検証することを目的とした。まず、両者の評価スコアの平均値に有意な差が存在するかを明らかにするため、各観点において対応のある  $t$  検定 (Paired-Samples  $t$ -Test) を実施し、その差の大きさを定量的に把握するために効果量 (Cohen's  $d$ ) を算出した。次に、AI 評価と教員評価の傾向の違いを視覚的に把握するため、各観点における「教員評価－AI 評価」の差分スコア

を算出し、箱ひげ図(boxplot)によってその分布を可視化することで、AIが特定の観点において一貫して過小または過大評価を行う傾向の有無を検証した。加えて、AIと教員がそれぞれどの程度安定した評価を行っているかを確認するため、英作文ごとに各観点スコアの標準偏差を算出し、評価のばらつきを比較した。さらに、両者の観点別スコアや平均値がどの程度連動して変化するかを確認するため、Pearson's  $r$  を用いて相関分析を行った。これらすべての統計分析はSPSSを用いて実施した。

### 3.4 AIの活用に対する意識調査

AIによる英作文評価に対する教員の受容意識を把握するため、評価実施の前後にアンケート調査を実施した。調査設計にあたっては、Davis(1989)のTechnology Acceptance Model(TAM)を理論的枠組みとして採用し、「知覚された有用性(Perceived Usefulness)」「知覚された使いやすさ(Perceived Ease of Use)」「行動意図(Behavioral Intention:BI)」「利用態度(Attitude Toward Using:ATT)」「社会的影響(Social Influence:SI)」の5構成要素に対応する計32項目を作成した。すべての設問は5段階のリッカート尺度により構成されており、回答者の主観的評価を数量化できるよう設計されている。事前アンケートでは、年齢層、指導年数、ICT活用頻度、AIの使用経験、AIへの期待感や抵抗感など、教員の基本属性および初期意識を収集した。事後アンケートでは、実際のAI評価体験を踏まえた印象や今後の導入意向を把握することを目的とした(資料1)。ただし、本研究では教員の一般的な受容傾向を明らかにすることに焦点を当てた。また、自由記述では生成AIを使用する際の懸念と期待を尋ね、TAMに基づく分析結果を考察する際の材料とした。

## 4 結果および考察

### 4.1 AI評価、教員評価の一致度分析(対応のあるサンプルの $t$ 検定)

英検準拠ルーブリック(内容、構成、語彙、文法)および新学習指導要領準拠ルーブリック(知識、表現、主体的態度)の各観点において、AI評価と教員評価との一致度を検証するため、対応のあるサンプルの $t$ 検定を実施した。さらに、差の大きさを定量的に把握するために効果量(Cohen's  $d$ )を算出し、その解釈には Plonsky and Oswald(2014)の基準を用いた。加えて、両者の関連性を確認するために Pearson's  $r$  を算出した。結果の概要を表1および表2に示す。

英検準拠ルーブリック(表1)については四つの観点のうち、「内容」に関しては AI と教員の評価の平均がほぼ同一であり、有意な差は確認されなかった。また、相関係数は非常に高く、AIによる判断が教員とおおむね一致していることが示された。一方、「構成」「語彙」「文法」においては、いずれも AI 評価の方が教員評価よりもやや低く、統計的に有意な差が認められた。特に「文法」では、効果量  $d = -0.58$  が中程度に該当し、AI が文法面を過小評価する傾向が示唆された。これに対し、「構成」「語彙」の効果量は 小程度にとどまり、実質的な差は限定的であった。ただし、これらの観点においても中～高程度の相関が確認され、評価の傾向は基本的に整合していたといえる。

学習指導要領準拠ルーブリック(表2)の3観点では、「知識」と「主体的態度」において、AI の評価が教員より有意に低いことが明らかとなった。特に「知識」の効果量  $d = -0.80$ は 大きい水準に該当し、AI と教員の認識の違いが顕著であった。また、「主体的態度」の効果量  $d = -0.67$ は 中程度に位置づけられ、この観点でも AI 評価の限界が浮き彫りとなった。一方、「表現」については平均値の差は小さく、有意な差は認められなかった。効果量  $d = -0.18$  は 小程度であり、実質的な差はほとんどないと解釈できる。これら3観点すべてにおいて、中～高程度の正の相関が確認され、AI 評価は全体として教員評価とよく一致していると考えられる。

以上のことから、AI評価と教員評価は観点ごとに一定の一貫性を示し、高い相関関係が確認された。しかし、英検準拠ルーブリックにおける「文法」や、新学習指導要領準拠ルーブリックにおける「知識」「主体的態度」では、AIが一貫して教員よりも厳しい評価を行う傾向が明らかとなった。これらの結果は、AI評価の有効性を支持する一方で、特定観点における過小評価の可能性にも留意する必要があることを示している。

■表1: 英検準拠ルーブリックにおけるAI評価と教員評価の比較 (N = 30)

| 観点 | 教員平均 (SD)  | AI 平均 (SD) | t 値   | p 値  | 効果量 (d) | 相関係数 (r) |
|----|------------|------------|-------|------|---------|----------|
| 内容 | 3.06 (.80) | 3.01 (.82) | -0.46 | .650 | -0.08   | .795     |
| 構成 | 2.97 (.70) | 2.76 (.81) | -2.19 | .037 | -0.40   | .765     |
| 語彙 | 3.05 (.72) | 2.85 (.76) | -2.31 | .028 | -0.42   | .722     |
| 文法 | 3.22 (.65) | 2.93 (.72) | -3.16 | .004 | -0.58   | .807     |

■表2: 学習指導要領準拠ルーブリックにおけるAI評価と教員評価の比較 (N = 30)

| 観点    | 教員平均 (SD)  | AI 平均 (SD) | t 値   | p 値   | 効果量 (d) | 相関係数 (r) |
|-------|------------|------------|-------|-------|---------|----------|
| 知識    | 2.11 (.55) | 1.80 (.52) | -4.38 | <.001 | -0.80   | .743     |
| 表現    | 2.01 (.48) | 1.93 (.51) | -1.00 | .325  | -0.18   | .716     |
| 主体的態度 | 2.12 (.50) | 1.85 (.53) | -3.66 | <.001 | -0.67   | .784     |

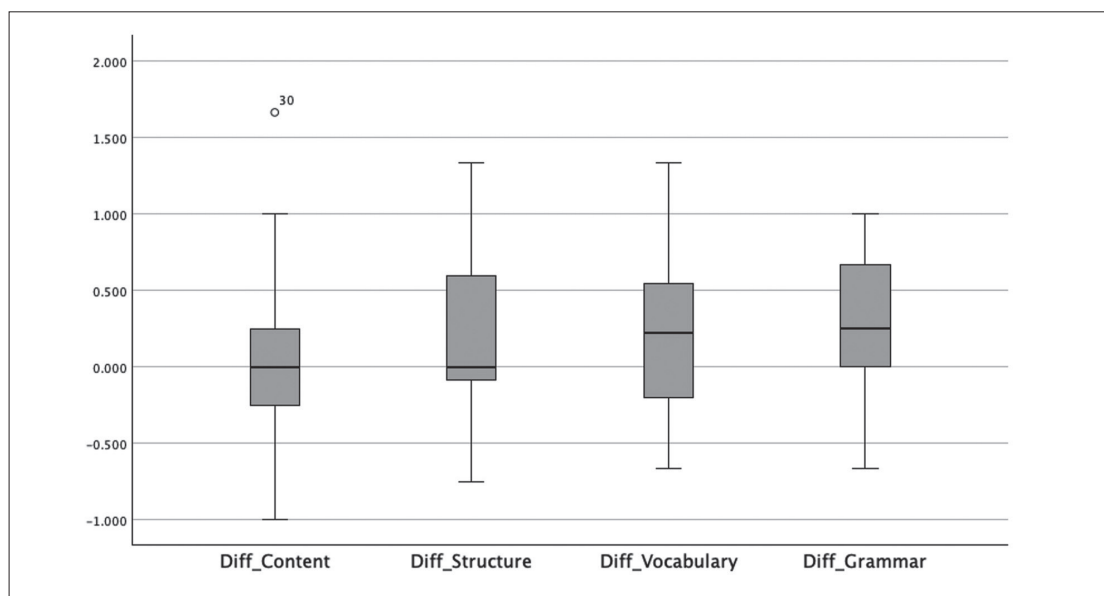
## 4.2 AI評価の傾向分析 (箱ひげ図)

AIと教員の評価傾向の違いを視覚的に捉えるため、各観点における「教員評価－AI評価」の差分スコアを算出し、その分布を箱ひげ図で示した。ここで正の値は教員評価の方がAI評価を上回っていること、負の値はAI評価が高かったことを意味する。本研究では、英検準拠の観点(内容、構成、語彙、文法)および学習指導要領準拠の観点(知識、表現、主体的態度)に分けて分析を行い、それぞれ図1および図2に示した。

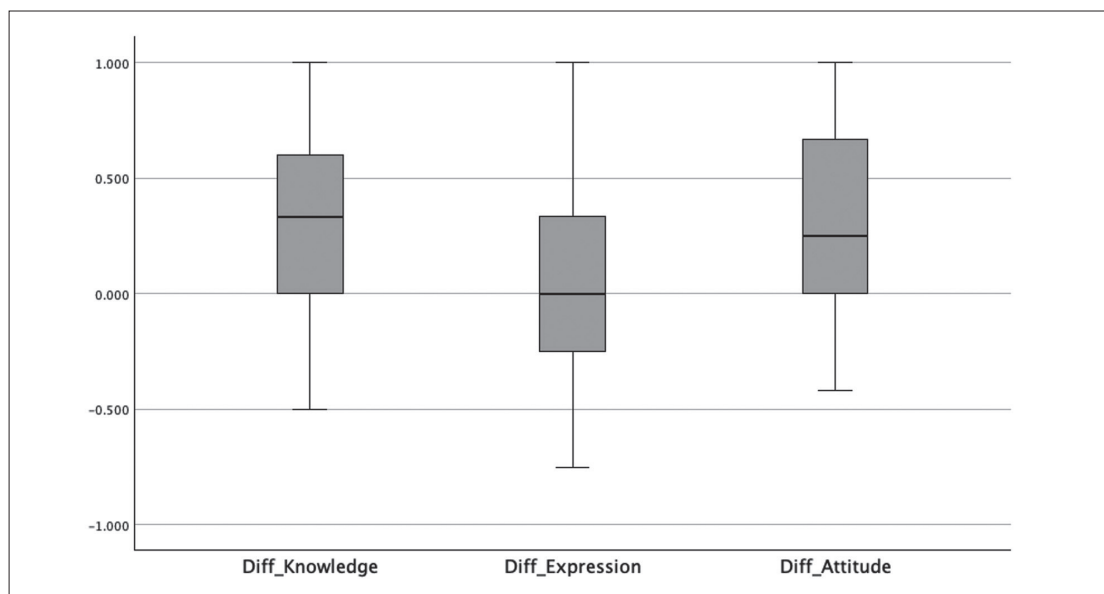
英検準拠ルーブリックについては(図1)、4観点すべてにおいてAI評価が教員評価を下回る傾向が明確に示された。特に「文法」においては、差分スコアの中央値が他の観点よりも大きく、AIがこの領域で一貫して厳しめの評価を行っていることがうかがえる。また、「語彙」や「構成」においても、教員評価の方がやや高い傾向が確認されており、いずれも評価差のばらつきは比較的少なく、AIの判断が一定の傾向に収束している様子が見て取れる。

学習指導要領準拠ルーブリックについては(図2)、特に「知識」および「主体的態度」において、教員評価がAI評価を明確に上回っており、差分の中央値はおおよそ +0.3～+0.5 の範囲に分布していた(図2)。内容(Content)の観点では作文30番に外れ値が確認され、AIが当該作文に著しく低いスコアを付与したことが示された。これは、AIが主観的な要素や学習態度のような内面的な観点を適切に判断することに限界がある可能性を示している。一方、「表現」については、中央値が0に近く、ばらつきも比較的小さかったことから、論理構成やアイデア展開といった比較的客観的な評価項目においては、AIと教員の判断が一致しやすいことがうかがえる。

以上の結果から、AIは「文法」や「構成」など明示的な形式的基準に対しては安定的かつ妥当な判断を下す一方、「知識」や「主体的態度」といった学習内容の正確さや学習への姿勢など、文脈や内面の理解が求められる観点では、AIが評価を低くつける傾向が見られた。特に「文法」「知識」「態度」に見られた評価差は、AI評価の限界を示唆するものであり、今後の活用には、観点ごとの特性やバイアスを考慮し、教員による補完的な判断との併用が望ましいといえる。



■ 図1: 英検準拠ルーブリックにおける教員評価 - AI評価の差 (箱ひげ図)



■ 図2: 新学習指導要領準拠ルーブリックにおける教員評価 - AI評価の差 (箱ひげ図)

### 4.3 評価の一貫性分析 (Intraclass Correlation Coefficient; ICC)

本研究では、教員および生成AIによる英作文評価の信頼性と一貫性を明らかにするため、各観点において級内相関係数 (Intraclass Correlation Coefficient: ICC) を用いて、評価の一致度を比較した。ICCは、同じ英作文に対して複数の評価者がどの程度似た評価をしているかを数値で示すもので、教育評価の分野でもよく使われている指標である (Shrout & Fleiss, 1979)。分析にあたっては、AIと教員の両方について、「Two-Way Random, Absolute Agreement, Average Measures」と呼ばれる共通の方法を用いてICCを算出した。これは、評価する人と英作文の両方をランダムな要素とみなして、どれくらい評価が一致しているかを確かめる方法である。AIも教員も同じ方法で分析したことで、どちらの評価がど

の観点で安定していたのかが明確になった(表3)。例えば、文法の観点ではAIの評価が非常に安定しており( $ICC = .919$ )、一方で「知識」や「主体的態度」など、内容の理解が必要な観点では、教員評価の方が一貫している傾向も見られた(表4)。 $ICC$ の数値の見方としては、Cicchetti(1994)の基準を参考にし、.75以上を「非常に良い(excellent)」, .60～.74を「良い(good)」, .40～.59を「中程度(fair)」, .40未満を「信頼性が低い(poor)」として分析を行った。

■表3: 英検準拠ルーブリックにおける教員評価とAI評価の $ICC$ 比較

| 観点 | 教員評価 $ICC$ | AI評価 $ICC$ | 評価傾向の比較                |
|----|------------|------------|------------------------|
| 内容 | .650       | .556       | 教員, AIともに fair 程度の一貫性  |
| 構成 | .747       | .885       | 教員:good, AI:excellent  |
| 語彙 | .915       | .880       | 両者とも excellent, 一貫性が高い |
| 文法 | .703       | .919       | 教員:good, AI:excellent  |

表3に示すとおり、語彙(Vocabulary)においては教員, AIいずれも非常に高い $ICC$ 値を示し、.75以上の「非常に良い(excellent)」水準に分類された。特に、両者とも.88を超える値を示しており、既存の基準における「excellent」の中でも際立って高い信頼性が確認された。このことから、語彙使用の適切さや多様性といった観点は、比較的客観的かつ明確な基準で評価されていることが示唆される。構成(Structure)および文法(Grammar)においても、AI評価はそれぞれ.885, .919と非常に高い $ICC$ を示し、教員評価(.747, .703)を上回った。これは、AIが構合力や文法的正確性に関して一貫した評価基準を保持していることを示している。一方、内容(Content)の観点では、教員評価が.650, AI評価が.556にとどまり、いずれも高い一貫性とは言えず、教員, AIともに fair レベルにとどまった。また、教員評価においては、95%信頼区間にマイナス値(CI: -0.210～0.959)が含まれており、統計的に一貫性が確認されたとは言いがたい結果であった。これは、主張の明確さや説得力といった定性的、主観的要素が強く、評価者間で判断が分かれやすい観点であることを示唆する。AIにおいても、意味解釈を要する観点では一貫性のある処理が難しい。

■表4: 学習指導要領準拠ルーブリックにおける教員評価とAI評価の $ICC$ 比較

| 観点    | 教員評価 $ICC$ | AI評価 $ICC$ | 評価傾向の比較               |
|-------|------------|------------|-----------------------|
| 知識    | .932       | .471       | 教員:excellent, AI:fair |
| 表現    | .721       | .900       | 教員:good, AI:excellent |
| 主体的態度 | .688       | .200       | 教員:good, AI:poor      |

学習指導要領準拠の観点においては(表4)、知識(Knowledge)で教員評価は.932と極めて高く、観定の定義や評価基準が教員間で共有されていたことが示される。一方で、AI評価は.471にとどまり、fair レベルであり、知識の適切さや情報の正確性の判断において一貫性が保たれていないことが示された。これは、文脈理解や背景知識の活用を必要とする観点において、生成AIの処理能力が限定的である可能性を示している。表現(Expression)については、AIが.900とexcellent レベルの高い一貫性を示し、教員評価(.721)よりも高い水準であった。これは、AIが論理構成、段落展開、接続表現などの形式的な構造要素に関して、安定した基準で評価をしていることを意味している。特に、構文や形式的特徴に強みを持つAIの特性が、この観点において顕著に発揮されたと考えられる。対照的に、主体的態度(Attitude)の観点では、教員評価でも.688とやや低めであり、AI評価は.200とpoorレベルに分類された。この結果は、学習への意欲や内面的な姿勢といった主観性の高い観点において、AIが一貫性をもって評価することが極めて難しいという実態を示している。

#### 4.4 教員の受容意識(TAM分析)

本研究では、生成AIを用いた英作文評価の有効性を検討するにあたり、その導入、活用に対する高等学校英語教員の受容意識を把握することを目的として、TAMに基づいた事前・事後アンケート調査を実施した(資料1)。調査対象は、本研究に協力した高等学校英語教員20名であり、TAMの五つの主要構成要素に対応する計32項目について、5段階リッカート尺度による評価を依頼した。構成要素は以下のとおりである:知覚された有用性(PU)、知覚された使いやすさ(PEOU)、行動意図(BI)、使用態度(ATT)、および社会的影響(SI)。質問項目の詳細は資料1「教員の受容意識(TAM分析)アンケート調査項目」に示した。分析においては各変数ごとに平均値を算出し、次に行動意図を目的変数とした重回帰分析(multiple regression analysis)を実施し、PU、PEOU、ATT、SIがBIにどのような影響を与えているかを検討した(表5)。分析の結果、「知覚された有用性」は「行動意図」に統計的に有意な負の影響を与えていた( $p < .05$ ) ( $\beta = -0.651$ )。これはTAMの理論的前提とは異なるものであった。本来、PUが高いほどBIも高くなる(=便利だと思えば使いたくなる)とされているが、今回のデータでは「便利だと感じても、使いたくない」という教員の認識が浮き彫りになった。この逆転現象の背景には、アンケートの自由記述で述べられた次のような意見が関係していると考えられる:

「AIで評価しても結局妥当性を確認するために読み直す必要があり、手間が増えるだけ」「プロンプトによって結果が大きく変わるの不安」「生徒のことをきちんと理解するには、やはり自分で英作文を読む必要がある」つまり、教員はAI評価の利便性や機能面を理解していても、それが現場での実践に直結しないと感じている可能性がある。評価に対する責任感や、生徒との教育的関係を重視する姿勢が、AIの導入を躊躇する要因となっているのかもしれない。また、「便利だがプロンプト次第で結果が変わる」という意見に見られるように、AIの安定性や妥当性に対する不安感が、知覚された有用性(PU)を逆方向に作用させているとも考えられる。

■表5: BI(行動意図)に対する各要因の重回帰分析結果( $n=20$ )

| 変数         | 標準化係数( $\beta$ ) | $t$ 値  | 有意確率( $p$ 値) | 有意性             |
|------------|------------------|--------|--------------|-----------------|
| 知覚された有用性   | -0.651           | -2.558 | .022         | 有意( $p < .05$ ) |
| 知覚された使いやすさ | 0.254            | 1.103  | .287         | 有意でない           |
| 利用態度       | 0.208            | 0.707  | .490         | 有意でない           |
| 社会的影響      | 0.320            | 1.161  | .264         | 有意でない           |

#### 4.5 考察

本研究では、生成AIによる高校生英作文のルーブリック評価の妥当性、AI評価が実際の教育実践にどのように受け入れられるのかを検証することを目的とした。以下では、本研究で設定した研究質問(RQ)に基づき、得られた結果を踏まえて考察を行う。

RQ1:生成AIによる英作文評価は、教員による評価とどの程度一致するのか。

➡観点によって一致度に差があることが明らかになった。

対応のある $t$ 検定および相関係数の結果、「文法(Grammar)」「語彙(Vocabulary)」「構成(Structure)」といった形式的・明示的な観点では、AIと教員の評価は平均値の差が小さく、相関係数も高かったことから、おおむね一致していることが示された。一方、学習指導要領準拠によるルーブリック評価における主体的態度(Attitude)では、AIが教員より低いスコアを付ける傾向があることが分かった。なお、評価の信頼性(安定性)を確認する目的で教員、AIそれぞれの内部評価の安定性を算出したICC(Intraclass Correlation

Coefficient)においては、AI評価では形式的観点では高いICCが得られた一方、AttitudeではICC = .200と低い値を示した。教員評価においても観点によりICCの差が見られたが、概ね中～高の安定性を示している。このことは、AIが学習者の内面の意図や学習姿勢などを評価することが困難であることを示しており、Xiao and Zhi(2023)やPolakova and Ivenz(2024)の報告と一致する。

RQ2: 評価観点ごとに、生成 AI に特有の過小または過大評価の傾向は見られるか。

➡ AI は一部の観点で一貫して教員より厳しい傾向を示した。

箱ひげ図による視覚化(図1, 図2)では、英検準拠の4観点(内容, 構成, 語彙, 文法)すべてにおいて、AIの方が教員評価よりもやや低いスコアを付ける傾向が確認された。特に文法においては、差分スコアの中央値が大きく教員評価優位に偏っており、AIの評価基準がより厳格である可能性がある。また、新学習指導要領に準拠した観点でも、知識, 態度においてAIは一貫して過小評価する傾向を示しており、特に主観的評価におけるAIの限界を浮き彫りにした。さらに、両評価の関連性をピアソンの積率相関係数によって検証したところ、英検準拠ルーブリックでは内容( $r = .790, p < .001$ ), 構成( $r = .756, p < .001$ ), 語彙( $r = .716, p < .001$ ), 文法( $r = .812, p < .001$ )と、いずれも高い相関が得られた。新学習指導要領準拠ルーブリックにおいても、知識( $r = .739, p < .001$ ), 表現( $r = .747, p < .001$ ), 態度( $r = .784, p < .001$ )と、中程度から高い正の相関が確認された。以上から、AI評価は観点ごとに教員評価と一定の一貫性を有する一方で、特定の観点においては厳格な評価傾向を示すことが明らかになった。

RQ3: RQ1と2で明らかになった一致度や傾向差を前提として、高等学校英語教員は生成 AI による英作文評価をどのように受容しているか。

➡ 「使える場面もあるが、主観的評価は教員が担うべき」という条件付きの受容が多く見られた。

TAMに基づくアンケート分析では、「知覚された有用性(PU)」と「使いやすさ(PEOU)」はいずれも平均3.5以上と高く評価され、AIツールとしてのポテンシャルは一定程度認識されていることが分かった。一方、「行動意図(BI)」は3.3とやや低めであり、回帰分析ではPUがBIに統計的に有意な負の影響( $\beta = -0.651, p = .022$ )を与える結果となった。これは、AIを「理論上は有用だと考えているが、実際の授業では積極的に使いたいとは限らない」というジレンマを反映している可能性がある。自由記述では、「AI評価は補助的に使用するのが現実的であり、主観的観点の判断や生徒理解には教員の介入が不可欠」とする声が多く、AIの過信を戒める姿勢が顕著だった。また、「プロンプトによって出力結果が変わるため、教員自身にプロンプト設計の知識が必要」といった指摘もあり、AI活用における教員の新たなリテラシーの重要性も示唆された。

以上の結果から、生成 AI による英作文評価は、形式的観点においては信頼性と一貫性を確保しうるものの、主観的, 意味的観点においては限界があり、全面的な代替は困難であることが分かった。したがって、AIはあくまで補助的な評価ツールとして活用し、評価の中心は引き続き教員が担う必要がある。また、AIの評価結果を学習者自身が振り返る材料として活用することで、深い学びを促進する可能性もある。今後は、「AI評価をどう授業設計に組み込むか」「プロンプト設計力をどう育成するか」が重要な課題となるだろう。

## 5

## 結論および今後の課題

本研究では、高等学校英語教育におけるAI活用の可能性と課題を明らかにすることを目的に、生成 AI による英作文評価に対して、教員との観点別比較, ICC分析, TAMアンケートを通じた意識調査を実施した。分析の結果、AIは文法, 語彙, 構成といった形式的観点において非常に高い評価の一貫性を示し、教育現場においてこれらの観点の一部をAIが代替する可能性が示唆された。一方で、意味内容や態度といった文脈的, 主観的判断を要する観点では、評価のばらつきや限界が確認され、AIによる単独評価には慎重

な対応が求められることも明らかになった。教員アンケートの結果からは、AI評価の有用性や効率性に対する期待と同時に、評価の正確性や教育的責任に対する慎重な姿勢がうかがえた。AIの導入が教員の業務負担を軽減する一方で、学習者理解や指導の質の担保は人間が担うべきだという認識が多くの教員に共有されていた。特に、プロンプト設計やAIの出力の取扱いに関する教員の理解とスキルが、AI評価の信頼性と教育効果を大きく左右する可能性が示されたことは重要である。

今後の課題としては、第一に、複数のAIモデルを用いた比較研究を通じて、モデルごとの評価傾向や性能の違いを明らかにすること。第二に、AIの効果的な利活用に必要なプロンプト設計技術の向上を図り、教員が適切にAIを活用できる体制を構築すること。第三に、AIによる評価が学習者の表現力や学習意欲、思考の深化といった側面に中長期的に与える影響を検証する実証研究を継続的に行うことが挙げられる。また、本研究では英検準拠および新学習指導要領準拠の二種類のルーブリックを作成し、それらを基にAIと教員による評価を行ったが、参加者が評価に迷う場面が少なからず存在した。Stanley (2021) が「ルーブリックに関する知識の中で最も重要なことは、そのルーブリックが良いものか悪いものかを見極める能力である」と指摘するように、評価の目的に応じて適切かつ妥当なルーブリックを設計できる技術が、今後のパフォーマンス評価に関連した生成AI評価研究や教育実践において不可欠である。

## 謝辞

本研究を進めるにあたり、関西大学の竹内理教授から多大なるご指導とご助言を賜りました。研究活動に行き詰まっていた私に対し、的確なご指摘と温かな励ましを繰り返しくださり、再び研究に向き合う力を与えてくださいました。研究の構想から分析手法に至るまで、先生のご助言を通じて多くを学び、本研究を進めるうえで大きな支えとなりました。心から感謝申し上げます。あわせて、本研究に英作文を提供して下さった室戸高校の生徒の皆さん、ならびにご多忙の中にもかかわらず、評価およびアンケート調査に取り組んで下さった英語教員の先生方にも、深く感謝申し上げます。皆様のご協力と率直なご意見は、本研究にとってかけがえのない支えとなり、貴重な示唆を与えてくださいました。さらに、本研究は公益財団法人 日本英語検定協会による「英検研究助成制度」の助成を受けて実施されたことをここに記し、厚く御礼申し上げます。

## 倫理的配慮に関する指針

本研究は、高等学校英語教員を対象にAIによる英作文評価の有効性を調査するものであり、研究目的、方法、個人情報の取扱い等を記載した説明文書を事前に配布し、文書による同意を得て実施した。評価対象の英作文は匿名化し、個人が特定されないよう配慮した。AI評価にはChatGPTを用い、出力内容は教育的目的に限って使用した。教員アンケートも無記名で実施し、統計的に処理・分析した。

## 引用文献

- Asli, N. F., Mohd Matore, M. E. E., & Md Yunus, M. (2024). Construct validity of primary trait writing rubrics based on assessment use argument (AUA) validation framework. *Heliyon*, 10(22), e40053-. <https://doi.org/10.1016/j.heliyon.2024.e40053>
- Brookhart, S. M. (2013). How to create and use rubrics for formative assessment and grading. Ascd.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. Longmans, Green.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Chocarro, R., Cortiñas, M., & Marcos-Matás, G. (2023). Teachers' attitudes towards chatbots in education: a technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics. *Educational Studies*, 49(2), 295-313. <https://doi.org/10.1080/03055698.2020.1850426>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- EIbyary, K., Shabara, R., & Boraie, D. (2024). The differential role of AI-operated WCF in L2 students' noticing of errors and its impact on writing scores. *Language Testing in Asia*, 14(1), 59-24. <https://doi.org/10.1186/s40468-024-00312-1>
- Ellis, R., Loewen, S., & Erlam, R. (2006). IMPLICIT AND EXPLICIT CORRECTIVE FEEDBACK AND THE ACQUISITION OF L2 GRAMMAR. *Studies in Second Language Acquisition*, 28(2), 339-368. <https://doi.org/10.1017/S0272263106060141>
- Fan, Y., Tan, S., & Lim, G. Y. W. (2024). EAP teacher feedback in the age of AI: Supporting first-year students in EFL disciplinary writing. *Australian Journal of Applied Linguistics (Online)*, 7(3), 1943-. <https://doi.org/10.29140/ajal.v7n3.1943>
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10(3), 161-184. [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X)
- Fuller, L. P., & Bixby, C. (2024). The Theoretical and Practical Implications of OpenAI System Rubric Assessment and Feedback on Higher Education Written Assignments. *American Journal of Educational Research*, 12(4), 147-158.
- Han, J., & Li, M. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. *System (Linköping)*, 126, 103502-. <https://doi.org/10.1016/j.system.2024.103502>
- Kaldaras, L., Yoshida, N. R., & Haudek, K. C. (2022). Rubric development for AI-enabled scoring of three-dimensional constructed-response assessment aligned to NGSS learning progression. *Frontiers in Education (Lausanne)*, 7. <https://doi.org/10.3389/educ.2022.983055>
- Kim, Haeun & Baghestani, Shireen & Yin, Shuhui & Karatay, Yasin & Kurt, Sebnem & Beck, Jeanne & Karatay, Leyla. (2024). ChatGPT for Writing Evaluation: Examining the Accuracy and Reliability of AI-Generated Scores Compared to Human Raters. 10.31274/isudp.2024.154.06.
- Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, 6, 100225.
- Korzynski, P., Mazurek, G., Krzykowski, P., & Kurasinski, A. (2023). Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT. *Entrepreneurial Business and Economics Review*, 11(3), 25-37.
- Lai, C. Y., Cheung, K. Y., & Chan, C. S. (2023). Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: An extension of the technology acceptance model. *Computers and Education. Artificial Intelligence*, 5, 100178-. <https://doi.org/10.1016/j.caeai.2023.100178>
- Li, J., Jangamreddy, N. K., Hisamoto, R., Bhansali, R., Dyda, A., Zaphir, L., & Glencross, M. (2024). AI-assisted marking: Functionality and limitations of ChatGPT in written assessment evaluation. *Australasian Journal of Educational Technology*. <https://doi.org/10.14742/ajet.9463>
- Lin, S., & Crosthwaite, P. (2024). The grass is not always greener: Teacher vs. GPT-assisted written corrective feedback. *System (Linköping)*, 127, 103529-. <https://doi.org/10.1016/j.system.2024.103529>
- 文部科学省. (2024). 初等中等教育段階における生成AIの利活用に関するガイドライン (Ver.2.0). [https://www.mext.go.jp/content/20241226-mxt\\_shuukyo02-000030823\\_001.pdf](https://www.mext.go.jp/content/20241226-mxt_shuukyo02-000030823_001.pdf)
- Mohammed, S. J., & Khalid, M. W. (2025). Under the world of AI-generated feedback on writing: mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia*, 15(1), 7-26. <https://doi.org/10.1186/s40468-025-00343-2>
- Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language learning*, 64(4), 878-912.
- Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2410101>
- Pozdniakov, S., Brazil, J., Abdi, S., Bakharia, A., Sadiq, S., Gašević, D., Denny, P., & Khosravi, H. (2024). Large language models meet user interfaces: The case of provisioning feedback. *Computers and Education. Artificial Intelligence*, 7, 100289-. <https://doi.org/10.1016/j.caeai.2024.100289>
- Shermis, M. D., Koch, C. M., Page, E. B., Keith, T. Z., & Harrington, S. (2002). Trait Ratings for Automated Essay Grading. *Educational and Psychological Measurement*, 62(1), 5-18. <https://doi.org/10.1177/0013164402062001001>

## 引用文献

- Shin, D., & Lee, J. H. (2024). Exploratory study on the potential of ChatGPT as a rater of second language writing. *Education and Information Technologies*, 29(18), 24735-24757. <https://doi.org/10.1007/s10639-024-12817-6>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Stanley, T. (2021). Using rubrics for performance-based assessment: A practical guide to evaluating student work. Routledge.
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, Article 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Tate, T. P., Steiss, J., Bailey, D., Graham, S., Moon, Y., Ritchie, D., Tseng, W., & Warschauer, M. (2024). Can AI provide useful holistic essay scoring? *Computers and Education. Artificial Intelligence*, 7, 100255-. <https://doi.org/10.1016/j.caeai.2024.100255>
- Teimouri, Y. (2017). L2 SELVES, EMOTIONS, AND MOTIVATED BEHAVIORS. *Studies in Second Language Acquisition*, 39(4), 681-709. <https://doi.org/10.1017/S0272263116000243>
- Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education. Artificial Intelligence*, 7, 100270-. <https://doi.org/10.1016/j.caeai.2024.100270>
- Tran, T. T. T. (2025). Enhancing EFL Writing Revision Practices: The Impact of AI- and Teacher-Generated Feedback and Their Sequences. *Education Sciences*, 15(2), 232-. <https://doi.org/10.3390/educsci15020232>
- Troia, G. (2014). Evidence-based practices for writing instruction (Document No. IC-5). Retrieved from University of Florida, Collaboration for Effective Educator, Development, Accountability, and Reform Center website: <http://ceedar.education.ufl.edu/tools/innovation-configuration/>
- Uyar, A. C., & Büyükhiska, D. (2025). Artificial intelligence as an automated essay scoring tool: A focus on ChatGPT. *International Journal of Assessment Tools in Education*, 12(1), 20-32. <https://doi.org/10.21449/ijate.1517994>
- Yang, Yang. (2024). The Reliability of using ChatGPT in Rating EFL Writings. *Shanlax International Journal of Education*. 12. 49-59. 10.34293/education.v12i4.7855.
- Yao, Y., Zhu, X., Xiao, L., & Lu, Q. (2025). Secondary school English teachers' application of artificial intelligence-guided chatbot in the provision of feedback on student writing: An activity theory perspective. *Journal of Second Language Writing*, 67, 101179-. <https://doi.org/10.1016/j.jslw.2025.101179>
- Yavuz, F., Çelik, Ö., & Yavaş Çelik, G. (2025). Utilizing large language models for EFL essay grading: An examination of reliability and validity in rubric-based assessments. *British Journal of Educational Technology*, 56(1), 150-166. <https://doi.org/10.1111/bjet.13494>
- VanPatten, B., & Williams, J. (2007). *Theories in second language acquisition*. NY: Routledge.
- Xiao, Y., & Zhi, Y. (2023). An Exploratory Study of EFL Learners' Use of ChatGPT for Language Learning Tasks: Experience and Perceptions. *Languages (Basel)*, 8(3), 212-. <https://doi.org/10.3390/languages8030212>
- Zeevy-Solovey, O. (2024). Comparing peer, ChatGPT, and teacher corrective feedback in EFL writing: Students' perceptions and preferences. *Technology in Language Teaching & Learning*, 6(3), 1482-. <https://doi.org/10.29140/tltl.v6n3.1482>

## 資料1: 教員の受容意識(TAM分析)アンケート調査項目

## TAM 事前アンケート (生成 AI に関する意識)

1. PU 生成 AI は, 英作文評価を効率的に思う。
2. PU 生成 AI を使うことで, 英作文評価精度が向上すると思う。
3. PU 生成 AI による評価は, 生徒の英作文へのフィードバックの質を高めると思う。
4. PU 生成 AI による評価は, 教員による英作文評価よりも公平だと思ふ。
5. ATT 英作文の評価に生成 AI を使用することは, 教育現場において価値があると思う。
6. ATT 生成 AI を使用することに対してよい印象を持っている。
7. ATT 生成 AI を評価に取り入れることは, 教育実践を改善する助けになると思う。
8. ATT 生成 AI を評価に使用することに対して抵抗感を感じる。
9. SI 同僚教員は生成 AI を使用することを推奨していると思う。
10. SI 生成 AI を使わないと周囲から取り残されると感じる。
11. SI 業務の質向上のために生成 AI を使うことが望ましいと思う。
12. SI 業務の効率化のために生成 AI を使うことが望ましいと思う。

## TAM 事後 (生成 AI 使用後の評価)

13. PU 生成 AI は, 英作文評価を効率的に思う。
14. PU 生成 AI を使うことで, 英作文評価精度が向上すると思う。
15. PU 生成 AI による評価は, 生徒の英作文へのフィードバックの質を高めると思う。
16. PU 生成 AI による評価は, 教員による英作文評価よりも公平だと思ふ。
17. PEOU 生成 AI の操作方法は簡単だと感じる。
18. PEOU 英作文評価のプロセスに生成 AI を組み込むのは簡単だと思ふ。
19. PEOU 生成 AI を使い始める際, 特別な知識やスキルは必要ないと思ふ。
20. PEOU 生成 AI が生成した結果を理解し, 活用するのは簡単だと思ふ。
21. PEOU 業務において, 生成 AI を日常的に使用することに対して疑問やストレスを感じる。
22. BI 英作文評価において, 生成 AI を積極的に使用したいと思ふ。
23. BI 生成 AI が評価ツールとして無料で利用可能な場合, すぐに試してみたいと思ふ。
24. BI 他の教員にも生成 AI の使用を勧めたいと思ふ。
25. ATT 英作文の評価に生成 AI を使用することは, 教育現場において価値があると思ふ。
26. ATT 生成 AI を使用することに対してよい印象を持っている。
27. ATT 生成 AI を評価に取り入れることは, 教育実践を改善する助けになると思ふ。
28. ATT 生成 AI を評価に使用することに対して抵抗感を感じる。
29. SI 同僚教員は生成 AI を使用することを推奨していると思ふ。
30. SI 生成 AI を使わないと周囲から取り残されると感じる。
31. SI 業務の質向上のために生成 AI を使うことが望ましいと思ふ。
32. SI 業務の効率化のために生成 AI を使うことが望ましいと思ふ。

資料2: 教員の受容意識(TAM分析)主要構成要素の集計結果.....

|     | PU   | PEOU | BI   | ATT  | SI   |
|-----|------|------|------|------|------|
| T1  | 5    | 3    | 1    | 4    | 3.75 |
| T2  | 3.78 | 4    | 3.33 | 3.75 | 2.88 |
| T3  | 4    | 4    | 4    | 3.75 | 3.5  |
| T4  | 4    | 4    | 3.33 | 3.75 | 3.88 |
| T5  | 3.78 | 3.75 | 3.67 | 4    | 4.13 |
| T6  | 4.67 | 4    | 3.33 | 4.13 | 3.75 |
| T7  | 4    | 2.5  | 3.67 | 3.38 | 3.13 |
| T8  | 4.89 | 4    | 3.67 | 4.13 | 3.63 |
| T9  | 4    | 4    | 3    | 3.5  | 2.88 |
| T10 | 4.11 | 3.75 | 3.33 | 3.5  | 3.75 |
| T11 | 3.89 | 4.5  | 3.33 | 3.5  | 3    |
| T12 | 4.56 | 4.5  | 3.67 | 4.38 | 4.25 |
| T13 | 4    | 3.75 | 3.33 | 4.38 | 3.88 |
| T14 | 3.67 | 3.25 | 3.33 | 3.13 | 2.38 |
| T15 | 3.67 | 3    | 3.67 | 3.38 | 3.5  |
| T16 | 3.89 | 3.5  | 3.33 | 4.13 | 2.88 |
| T17 | 4    | 3.25 | 4.33 | 4.38 | 4.13 |
| T18 | 3.44 | 3    | 2.67 | 3.13 | 2.75 |
| T19 | 3.78 | 3.5  | 3.33 | 4.13 | 2.75 |
| T20 | 4.33 | 4.75 | 3.67 | 4.5  | 3.75 |

## 第37回 研究助成

## A 研究部門・報告Ⅳ・英語能力テストに関する研究

# 基本的心理欲求の充足が小学校における 外国語学習の楽しさと学習者エンゲージメント に与える影響

研究者: 中西 洋平 大阪府／関西大学大学院 在籍

《研究助言者: 寺内 一》

## 概要

2020年度の小学校学習指導要領改訂により、小学校5・6年生に外国語科が導入された。このような改訂の背景、また、第二言語習得分野において児童を対象とした実証研究の数に限りがある現状を踏まえ、本研究では、日本の小学校5・6年生の計369名を対象に、基本的心理欲求の充足が外国語学習の楽しさを媒介してエンゲージメントに及ぼす影響を検討した。分析には、理論検証を目的とした共分散構造方程式モデリングではなく、モデルの予測に重点を置く部分最小二乗構造方程式モデリング(Partial Least Squares Structural Equation Modeling)を用いた。得られた結果より、基本的心理欲求の充足が外国語学習の楽しさの三つの因子それぞれに有意な影響を与え、さらに外国語学習の楽しさがエンゲージメントに影響を及ぼすことが示された。一方で、モデルに含まれる基本的心理欲求、外国語学習の楽しさ、エンゲージメントの各要因については弁別的妥当性が十分に担保されず、児童を対象に複雑なモデルを検証することの難しさも明らかとなった。

## 1 はじめに

第二言語習得分野においては、これまで学習者の動機づけをはじめとする心理的要因に関する多くの研究が行われてきた(Boo et al., 2015)。また、学習者の感情面についても、外国語学習不安などを中心に、これまで多くの実証研究が実施されてきた(Teimouri et al., 2023; Zhang, 2019)。近年では、ポジティブ心理学や教育心理学の知見が第二言語習得分野にも取り入れられており、外国語学習不安などのネガティブな側面だけでなく、学習者が学びの過程で経験する様々な感情を包括的に捉えている。そのため、学習場面における達成感や学びへの興味・関心といった「外国語学習の楽しさ(Foreign Language Enjoyment)」などのポジティブな感情に関する研究にも注目が集まっている(Botes et al., 2022; Dewaele & MacIntyre, 2014, 2016; MacIntyre & Vincze, 2017; Shirvan et al., 2024)。さらに、従来の動機づけ理論だけでは十分に捉えきれない学習者の実際の行動や参加度に注目し、学習を多面的・包括的に捉える「エンゲージメント(Engagement)」という概念も注目を集めている(Hiver et al., 2024; Teravainen-Goff, 2022, 2023; Vo, 2024; Zhou et al., 2021)。エンゲージメントは、主に認知的・感情的・行動的エンゲージメントから構成されており(Fredricks et al., 2004)、学びのプロセスを包括的に捉え、学習に与える影響を裏づける実証研究の数も増えつつある(e.g., Eerdemutu et al., 2024; Khajavy, 2021)。

上述のような「外国語学習の楽しさ」や「エンゲージメント」に関する研究は、学習者の心的要因や実際の学習行動を理解する上で重要な役割を果たしているが、主な研究対象は大学生などの成人であり、本研究が対象とする児童を扱った研究は少ないのが現状である(Boo et al., 2015; Hiver et al., 2024;

Sudina, 2021)。例えば、Sudina(2021)は、2009年から2019年にかけて、外国語学習不安と動機づけに関する113の実証研究をレビューしたが、12歳以下の参加者を対象とした研究はわずか3件であったと報告している。同様に、Hiver et al.(2024)は、20年間のエンゲージメント研究をレビューした結果、7歳から12歳を対象とした研究は、112本中8本であったと報告している。これらのレビュー研究が示すように、こうした研究における参加者の偏りは第二言語習得分野全体で大きな課題であると指摘されている(Andringa & Godfroid, 2020)。

上述の第二言語習得分野における児童を対象とした研究の少なさに加えて、2020年度の小学校学習指導要領改訂により、外国語科が5・6年生に導入されたことは大きな転換点となっている。これまで週1回の外国語活動として行われていた学習が週2回の授業となり、検定教科書を用いた指導や教科としての評価が導入されたことで、児童の学習経験は大きく変わってきていると考えられる。

そこで、本研究では、第二言語習得分野において児童を対象とした研究数が限られている現状(Boo et al., 2015; Hiver et al., 2024; Sudina, 2021)、そして、2020年度の学習指導要領改訂により小学校に外国語科が正式に導入されたという背景を踏まえ(文部科学省, 2017)、外国語科における児童の教室における学習体験をよりよく理解するため、学習者の感情や動機づけといった心理的要因に着目する。特に、児童にとって、外国語学習の楽しさが学習に大きな影響を与えることから(Botes et al., 2022)、その前提条件となる基本的心理欲求の充足が外国語学習の楽しさに与える影響、そして、外国語学習の楽しさがエンゲージメントに与える影響を明らかにすることを目的とする。

## 2 先行研究

本研究では、三つの重要な概念である「外国語学習の楽しさ」、「エンゲージメント」、「基本的心理欲求」を扱う。ここでは、それぞれの理論的背景、定義、測定方法、および、これらの概念を用いた実証研究、ならびにそれらの関係性について検討した先行研究をまとめる。

### 2.1 外国語学習の楽しさ

近年、ポジティブ心理学の発展とその影響を受けて、第二言語習得分野でも研究の視点が大きく変わりつつある。これまでは外国語学習の際に生じる不安に注目した研究が多かったが(Teimouri et al., 2023; Zhang, 2019)、現在では外国語学習の過程で感じる達成感や楽しさといったポジティブな感情にも関心が向けられつつある(MacIntyre & Gregersen, 2012)。このような背景をもとに、第二言語習得分野で、先駆的な研究であるDewaele and MacIntyre(2014)は、「外国語学習の楽しさ(Foreign Language Enjoyment)」を概念化した。この概念の中心となるEnjoymentについて、彼らは、フロー理論(Csikszentmihalyi, 2008)を援用し、“Enjoyment is indicative of a state in which psychological needs are being met”(p. 242)と定義している。これは、外国語学習の楽しさが、後述する基本的心理欲求の充足と密接に関連していることを示している。

外国語学習の楽しさを研究するにあたり、Dewaele and MacIntyre(2014)の研究では、21項目からなる「外国語学習の楽しさ尺度(Foreign Language Enjoyment Scale)」を用いている。その後、Dewaele and MacIntyre(2016)は、この21項目からなる外国語学習の楽しさ尺度を使用して収集したデータに対して探索的因子分析を行い、「外国語学習の楽しさ」を構成する二つの因子を明らかにしている。一つ目は、「外国語学習の楽しさ(社会的側面)」であり、クラスメートとの関係、教師との関係、教室の雰囲気の良いなどが含まれる。二つ目は、「外国語学習の楽しさ(個人的側面)」であり、達成感や学習の充実感などによって構成されている。

第二言語習得分野において、「外国語学習の楽しさ」に関する研究は、この二つの研究(Dewaele &

MacIntyre, 2014, 2016) がきっかけとなり、この10年間で様々な研究が行われるようになった。例えば、外国語学習の楽しさが他の学習者要因とどのような関係があるのか(e.g., Tsang & Dewaele, 2023)、測定方法を改良する研究(e.g., Li et al., 2018)が増えており、研究の幅が広がっている。Tsang and Dewaele(2023)は、英語を外国語として学ぶ香港の小学校3・4年生111名を対象に、外国語学習不安、外国語学習の退屈さ、外国語学習の楽しさが、エンゲージメントおよび英語運用能力(proficiency)に与える影響を調査した。分析の結果、外国語学習の楽しさがエンゲージメントを有意に予測していることが明らかとなった( $\beta = 0.78, p < .001$ )。また、英語運用能力に対しても、外国語学習の楽しさは有意に予測し( $\beta = 0.35, p = .03$ )、この研究では、外国語学習の楽しさのみがエンゲージメントと英語運用能力を予測する結果となった。外国語学習の楽しさを測定する尺度については、Dewaele and MacIntyre (2014)の21項目からなる尺度やDewaele and MacIntyre(2016)の14項目尺度をもとに、世界各地でその妥当性および信頼性の検証が行われている。例えば、Li et al.(2018)は中国の高校生を対象に、14項目の外国語学習の楽しさ尺度(Dewaele & MacIntyre, 2016)を中国語版に改訂し、その妥当性および信頼性を確認した。分析の結果、項目数は14項目から11項目に絞り込まれ、三つの因子(個人的楽しさ、教師関連の楽しさ、雰囲気に関する楽しさ)からなる中国版外国語学習の楽しさ尺度を作成した。Botes et al.(2021)は、Dewaele and MacIntyre(2014)の研究で使用したデータを再分析し、三つの因子(教師に対する肯定的評価、個人的楽しさ、社会的楽しさ)からなる9項目の短縮版外国語学習の楽しさ尺度を作成した。日本の文脈においては、Nakanishi and Takeuchi(2024)が、小学校5・6年生510名を対象に、Dewaele and MacIntyre(2014)の21項目から構成される外国語学習の楽しさ尺度を児童の発達段階に合わせて文言等を改訂し、3因子15項目からなる小学生版外国語学習の楽しさ尺度を開発した。このように、ポジティブ心理学の知見を取り入れながら、外国語学習におけるポジティブな感情の役割が注目を集めており、そのための測定尺度の開発や、実証研究によるその効果の検証を試みる研究も増加している。

## 2.2 エンゲージメント

第二言語習得分野において、学習者の動機づけに関する研究は長年にわたって積み重ねられてきた(Boo et al., 2015)。しかし、動機づけという心理的な概念だけでは捉えきれない、学習者の実際の行動や学習過程における具体的な取り組みにも注目する必要性が指摘されるようになった(Teravainen-Goff, 2022)。このような背景から、近年、主に教育心理学分野で研究されてきた学習者の行動に注目したエンゲージメントという概念(Fredricks et al., 2004)が、第二言語習得分野でも重要な理論的枠組みとして注目を集めている(Hiver et al., 2024)。エンゲージメントは多面的であり、学習環境や学習活動に応じて変化する動的な概念である(Sang & Hiver, 2021)。エンゲージメントの構成要素については様々な議論があるものの、多くの研究では観察可能な行動を中心に据え、行動的・認知的・情意的エンゲージメントの三要素を核として捉えている(Fredricks et al., 2004)。行動的エンゲージメントとは、課題への実際の取り組み時間や、ペアやグループ活動における能動的な取り組みなど、観察可能な行動として現れるものである。しかし、行動そのものに認知的・情意的要素が密接に関わっていることから、厳密に分けることの困難さも指摘されている(Sang & Hiver, 2021)。認知的エンゲージメントについて、Sang and Hiver(2021)は学習プロセスにおける認知的努力を指し、例えば、共同学習の場面で、相手の様子をうかがいながら説明を加えたり、意見を整理して伝えたりすることが挙げられる。情意的エンゲージメントとは、楽しさなどのポジティブな感情や、退屈といったネガティブな感情の表出として捉えることができる(Sang & Hiver, 2021)。

このように、エンゲージメントは多面的な側面から学習者の学びのプロセスを包括的に捉えるための有用な枠組みである一方で、その構成概念の妥当性については現在も議論が続いている(Hiver et al., 2024; Martin et al., 2017; Teravainen-Goff, 2023; Vo, 2024)。具体的には、各構成要素の間に重なりが見られる点や、動機づけの概念とエンゲージメントとの違いが曖昧である点が指摘されている。これらの

課題を受けて、エンゲージメントの概念を整理する取り組みや、エンゲージメントを測定するための質問紙の改善を図る研究が進みつつある。例えば、Vo(2024)は、これまでのエンゲージメントの定義や構成概念が曖昧であることを指摘した上で、ベトナムの大学生413人を対象に収集したデータに対して、探索的構造方程式モデリング(exploratory structural equation modeling)という高度な分析手法を用いて、使用したエンゲージメント尺度の妥当性および信頼性を検証した。そうした動きの中で、Teravainen-Goff(2023)は、これまでのエンゲージメント研究の文献を整理した上で、エンゲージメントの中心的定義である行動的側面を重視し、“engagement is an ultimately behavioural concept with underlying cognitive and affective facets”(p. 2)として、新たなエンゲージメント尺度の作成を試みた。質問項目の作成にあたっては、先行研究で使用されていた項目の活用、教員や学習者へのインタビュー調査結果からの項目作成、4名の言語教育専門家によるエキスパートチェックなどを経て、最終的に36項目を準備した。この質問項目をイギリスの中学生378人を対象に回答を依頼し、探索的因子分析および確証的因子分析を行った結果、五つの因子(1. 学習努力の強さ(Intensity of effort in learning), 2. 社会的エンゲージメントの強さ(Intensity of social engagement), 3. 教師とのエンゲージメントの質(Perceived quality of engagement with the teacher), 4. 仲間とのエンゲージメントの質(Perceived quality of engagement with peers), 5. 学習活動とのエンゲージメントの質(Perceived quality of engagement with learning activities))から構成される「第二言語学習エンゲージメントの強さと質に関する質問紙(the Intensity and Perceived Quality of L2 Engagement Questionnaire)」を完成させた。

エンゲージメントを取り入れた実証研究として、Jin(2024)は外国語学習の楽しさがエンゲージメントに及ぼす影響について、564名の中国人学習者を対象に調査を行い、外国語学習の楽しさからエンゲージメントを $\beta = 0.55$ で予測できることを示している。また、Hosseini et al.(2022)は、イランのEFL学習者386名を対象に同様の関係性を検証し、 $\beta = 0.43$ という結果を報告している。

## 2.3 自己決定理論と基本的心理欲求

自己決定理論とは、心理学や教育学などの幅広い分野で用いられている人間の動機づけのメカニズムを説明する理論的枠組みである。第二言語習得分野においては、Noels et al.(1999)の先駆的研究以降、20年以上にわたって実証研究が蓄積されている(Alamer et al., 2025; Al-Hoorie et al., 2025; Sugita-McEown & Oga-Baldwin, 2019)。自己決定理論は、人間の動機づけを多角的に説明する複数の理論から構成されており、第二言語習得分野においては、内発的動機づけと外発的動機づけの概念を中心とする有機的統合理論や基本的心理欲求理論などが広く用いられている。本研究が援用する基本的心理欲求理論は、三つの要素から構成されている(Sugita-McEown & Oga-Baldwin, 2019)。一つ目は、自律性(autonomy)である。これは、学習者が学びを自分で選択して、自分にとって価値のある活動として捉えているかということである。二つ目は、有能性(competence)である。これは、学習者が課題や活動に対して自信を持って取り組むことができるかに関連している。三つ目は、関係性(relatedness)である。これは、学習者がクラスメートや教師と良好な関係を築きながら学習に取り組んでいるかを示している。これら三つの基本的心理的欲求を満たすことが言語学習の動機づけにおいて重要であるとされている(Alrabai & Algazzaz, 2025)。

この基本的心理欲求理論を取り入れ、外国語学習の楽しさやエンゲージメントとの関係性に焦点を当てた実証研究では、Alrabai and Algazzaz(2025)がサウジアラビアの121名の大学生を対象にデータを収集し、学習者の基本的心理欲求の充足が外国語学習の楽しさを有意に予測し、さらにその楽しさが情意的エンゲージメントを高めることを示した。これは、基本的心理欲求の充足が外国語学習の楽しさを介して情意的エンゲージメントに影響を与えるという媒介プロセスを実証的に裏づけるものである。また、Wang(2024)は中国の大学生391名を対象とした調査において、基本的心理欲求の充足がポジティブな学習感情に及ぼす影響を明らかにした( $\beta = 0.620, p < 0.001$ )。

### 3 研究目的

上述の先行研究を踏まえて、本研究では、基本的心理欲求の充足が外国語学習の楽しさを媒介としてエンゲージメントに与える影響を部分最小二乗構造方程式モデリング (partial least squares structural equation modeling) によって明らかにする。この目的のために、以下の二つの研究課題を設定した。

#### 研究課題1

基本的心理欲求・外国語学習の楽しさ・エンゲージメントの測定モデルは、どの程度、信頼性および妥当性が確保されているか。

#### 研究課題2

基本的心理欲求の充足が外国語学習の楽しさを媒介としてエンゲージメントに与える影響はどの程度か。

### 4 研究方法

本研究では、調査対象である児童の発達段階を考慮し、すでに妥当性および信頼性が検証されている5件法の3尺度を用いた。各尺度の詳細については後述し、質問紙の項目は資料1を参照されたい。

#### 4.1 外国語学習の楽しさ尺度

本研究では、日本の小学校において妥当性および信頼性が検証された児童版外国語学習の楽しさ尺度を使用した (Nakanishi & Takeuchi, 2024)。この尺度は、Dewaele and MacIntyre (2014) による21項目の外国語学習の楽しさ尺度を基とし、まず成人版日本語学習の楽しさ尺度として妥当性および信頼性を検証し、その後児童版へと改訂されたものである。本尺度は15項目 (資料1, 1-15) で構成されており、三つの因子により外国語学習の楽しさを測定している。一つ目に、個人的要因を中心とする「外国語学習の楽しさ (個人的側面)」であり、具体的な質問項目として、「私は英語の授業では、自分ができたと感じる事が多い」、「私は英語の授業では、興味深いこと (もっと知りたいと思えること) を学んでいる」などの項目を含む。二つ目は、社会的要因を中心とする「外国語学習の楽しさ (社会的側面)」である。具体的な質問項目は、「英語の授業は、良い雰囲気だ」、「私たちは、英語の授業中によく笑い合う」などで構成される。三つ目は、教師要因を中心とする「外国語学習の楽しさ (教師的側面)」である。質問項目は、「英語の先生は、フレンドリー (親しみやすい) に接してくれる」、「英語の先生はいつも励まそうとしてくれる」などを含んでいる。各因子の信頼性係数は、尺度全体で  $\omega = 0.958$ 、外国語学習の楽しさ (個人的側面) = 0.897、外国語学習の楽しさ (社会的側面) = 0.911、外国語学習の楽しさ (教師的側面) = 0.914 をそれぞれ示している。

#### 4.2 エンゲージメント尺度

本研究では、Teravainen-Goff (2023) が作成した Intensity and Perceived Quality of L2 Engagement Questionnaire を土台として、成人版日本語尺度として改訂し、妥当性および信頼性を検証した後、児童版にさらに改訂して妥当性および信頼性を検証した尺度を使用した。この質問紙は23項目 (資料1, 28-50) からなる四つの因子によって構成されている。一つ目は、仲間との関わり合いの質 (perceived quality of engagement with peers) であり、具体的な質問項目として、「私はグループ活動 (班活動) の時、メンバー

と一緒にいて居心地が良いと感じる」,「私は普段から、クラスメートと一緒に活動することで、多くのことを学べていると感じている」などの項目を含む。二つ目は、教師との関わり方の質(perceived quality of engagement with teachers)であり、質問項目として、「英語の先生は普段から、私の英語学習がうまくいくように助けてくれる」,「英語の先生は普段から、英語学習に役立つヒントや手助けを与えてくれる」などで構成される。三つ目は、学習への努力の強さ(intensity of effort in learning)であり、「私は普段から、授業中に行われる活動に集中している」,「私は普段から、授業中の学習や活動には全力で取り組む」などを項目に含む。四つ目は、教材内容への取り組みの質(perceived quality of engagement with teaching content)であり、質問項目として、「授業で扱われる教材等(教科書など)は、英語を学ぶ上で役に立つ」,「授業の内容は、私の学びにつながると感じる」などから構成される。各因子の信頼性係数は、尺度全体で $\omega = 0.954$ , 各因子それぞれ0.917, 0.888, 0.87, 0.892であり、全体および各因子ともに高い信頼性を示している。

### 4.3 基本的心理欲求尺度

基本的心理欲求については、日本において成人版日本語基本的心理欲求尺度の妥当性および信頼性が既に検証されているため(Agawa & Takeuchi, 2016), この尺度を児童版に改訂して使用した。改訂にあたっては、前述の外国語学習の楽しさ尺度およびエンゲージメント尺度と同様に、エキスパートチェックを実施するなど厳密な手続きを踏んだ。本尺度は計12項目(資料1. 16-27)で構成され、三つの因子を測定している。一つ目は、自律性であり、「英語の授業では、先生は私たちの意見や発言をよく聞いてくれる」,「英語の授業で、先生は授業中に行う活動やその内容の大切さを説明してくれる」などの質問項目を含む。二つ目は、有能性である。具体的な質問項目として、「英語の授業では、自分の努力してきたことがうまくいったと感じることがある」,「英語の授業では、『できた』という気持ちになることがある」などで構成される。三つ目は、関係性であり、「英語の授業では、楽しくて心地よい雰囲気があると思う」,「英語の授業では、クラスの仲間と協力してやっていると思う」などの項目を含んでいる。本尺度の信頼性係数は、全体で $\omega = 0.969$ と高い値を示している。

### 4.4 参加者

本研究では、便宜的サンプリング法を採用した。データは、2025年1月から3月の期間に収集した。調査は3名の外国語専科教員の協力を得て実施された。なお、透明性の観点から、協力を得た教員はエンゲージメントなどの概念について理解しており、エンゲージメントを高めるような視点を重視した授業実践を行っている。本研究の調査対象は、学習指導要領に基づき週2回の外国語科授業を受けている公立小学校3校に在籍する小学5・6年生369名である。参加への同意が得られなかった11名を除外し、358名分のデータを分析対象とした。参加者の詳細に関して、男性169名、女性173名、回答しないが16名であった。学年別では5年生201名、6年生157名であった。年齢は、10歳15名、11歳200名、12歳143名であった。

### 4.5 分析方法

データ分析には、R(version 4.5.1)を使用した。分析の前には、児童の回答パターンの確認を行い、すべての質問紙項目に対して同一の選択肢を選択している回答者(例:すべての質問紙項目に対して5を選択)については、調査への注意深い取り組みが不十分である可能性が考えられるため、分析から除外することとした。その後、データの全体傾向を把握するため、各変数の記述統計量を算出した。

本研究では、研究課題を検証するために、理論検証を目的とした共分散構造方程式モデリングではなく、予測に重きを置く部分最小二乗構造方程式モデリングを採用した。その選択理由は二つある。一つ目

に、本研究は児童を対象としており、当該分野の研究の蓄積がまだ少ないため、理論検証よりも予測に重点を置く部分最小二乗構造方程式モデリングがより適していると考えた(Hair et al., 2022)。二つ目に、部分最小二乗構造方程式モデリングは、データが正規分布に従わなくても使用できるノンパラメトリックな分析方法であり(Hair et al., 2022)、図1(次頁)が示す通り、データの分布に偏りがあるため、この分析方法が適していると考えた。

分析については、Hair and Alamer(2022)の部分最小二乗構造方程式モデリングのガイドラインを参照した。分析にあたっては、SEMinRパッケージを用いた。SEMinRはオープンソースパッケージであり、部分最小二乗構造方程式モデリングの主要な機能をすべて実行することができる。本研究では、すべての構成概念を反動的(reflective)測定モデルとして設定し、測定モデルが適切かどうかを四つの基準で評価した。まず、項目信頼性については、各質問項目がどの程度その概念を測定できているかを因子負荷量で確認し、基準を0.7以上とした(Hair et al., 2022)。ただし、0.4から0.7の値でも、他の指標が良好であれば許容した(Hair & Alamer, 2022)。次に、内的一貫性については、各概念を測定する質問項目群の一貫性をCronbachの $\alpha$ 係数および合成信頼性(CR)で確認し、基準を0.7以上とした。収束的妥当性については、Average Variance Extracted (AVE)で確認し、基準を0.5以上とした。最後に、弁別的妥当性については、異なる概念が適切に区別されているかをHeterotrait-Monotrait Ratio(HTMT)で確認し、0.9未満を基準とした。

構造モデルについては、パス係数( $\beta$ )をブートストラップ(10,000回)で算出した。説明力は決定係数( $R^2$ )で評価した。さらに、外的妥当性はPLSpredictにより検証し、実際の観測値を反映した部分最小二乗構造方程式モデリング(PLS)と単純線形回帰モデル(LM)の二乗平均平方根誤差(RMSE)を比較した。詳細はHair and Alamer(2022)を参照されたい。

## 5 結果

データ分析を行う前に、全児童の回答パターンを確認し、すべての質問項目に対して同じ回答を選んだ参加者38名を分析対象から除外した。内訳は、全項目で「1」を選択した者が3名、「2」を選択した者が1名、「3」を選択した者が3名、「5」を選択した者が31名であった。これらを除いた残り320名のデータを用いて本分析を実施した。

### 5.1 記述統計

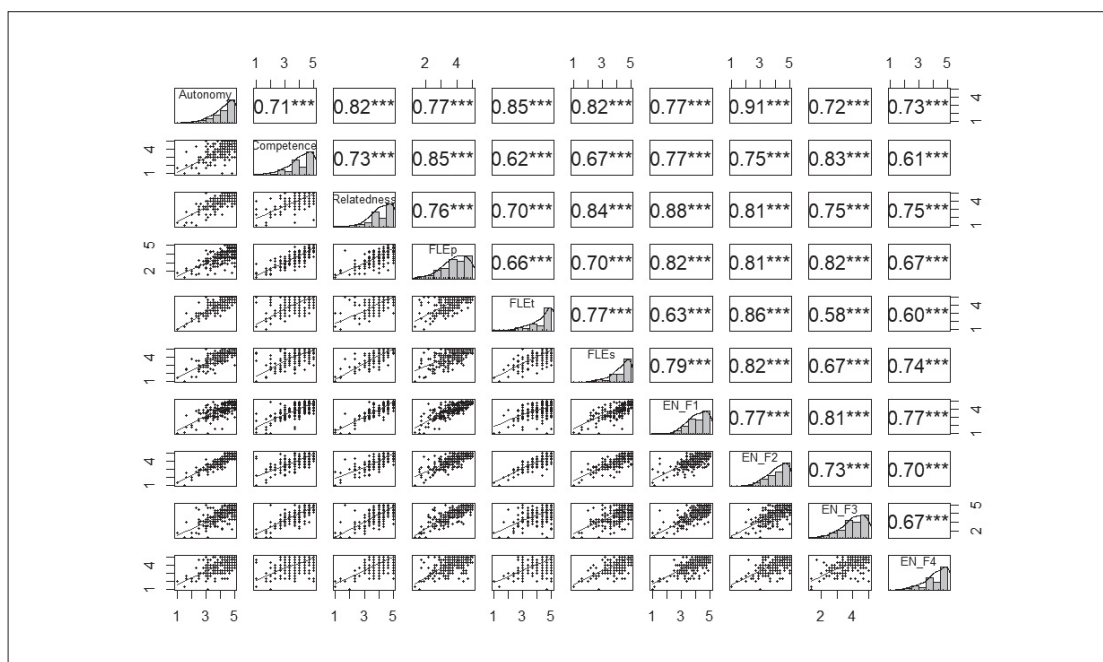
本研究で扱っている変数の記述統計結果を表1にまとめた。基本的心理欲求では、自律性が最も高く、有能感が最も低かった。外国語学習の楽しさについては、外国語学習の楽しさ(教師的側面)が最も高く、外国語学習の楽しさ(個人的側面)が最も低かった。学習者エンゲージメントの4因子は4.04から4.21の範囲にあった。標準偏差は0.76から0.88の範囲であった。すべての変数で歪度が負の値となっており、高得点側に偏った分布を示している。

■表1: 各変数の記述統計

| 変数                     | 平均値  | 標準偏差 | 歪度    | 尖度   |
|------------------------|------|------|-------|------|
| 基本的心理欲求(自律性)           | 4.26 | 0.77 | -1.24 | 1.44 |
| 基本的心理欲求(有能性)           | 4.08 | 0.88 | -1.15 | 1.16 |
| 基本的心理欲求(関係性)           | 4.19 | 0.80 | -1.26 | 1.98 |
| 外国語学習の楽しさ(個人的側面)       | 3.90 | 0.81 | -0.71 | 0.04 |
| 外国語学習の楽しさ(教師的側面)       | 4.36 | 0.79 | -1.35 | 1.53 |
| 外国語学習の楽しさ(社会的側面)       | 4.26 | 0.77 | -1.39 | 1.82 |
| エンゲージメント(仲間との関わりの質)    | 4.08 | 0.80 | -1.17 | 1.71 |
| エンゲージメント(教師との関わりの質)    | 4.16 | 0.76 | -1.04 | 1.06 |
| エンゲージメント(学習への努力の強さ)    | 4.04 | 0.79 | -0.97 | 0.71 |
| エンゲージメント(教材内容への取り組みの質) | 4.21 | 0.81 | -1.10 | 0.83 |

N = 320

図1は、基本的心理欲求、外国語学習の楽しさ、エンゲージメントの各次元の相関と分布を示したものである。ヒストグラムから、データの分布は右側に偏っており、正規分布から外れていることが分かる。基本的心理欲求の三つの要素(自律性、有能性、関係性)の間には、 $r = .71$ から $.82$ の強い正の相関が見られた。外国語学習の楽しさの各要素の間でも、 $r = .66$ から $.77$ の有意な正の相関が示された。エンゲージメントの四つの要素の間でも、 $r = .67$ から $.81$ の正の相関が確認された。また、基本的心理的欲求、外国語学習の楽しさ、エンゲージメントの各変数の間でも、 $r = .58$ から $r = .91$ の相関が認められた。



■図1: 基本的心理欲求・外国語学習の楽しさ・エンゲージメントの各次元の相関と分布

Note. Autonomy = 自律性, Competence = 有能性, Relatedness = 関係性, FLEp = 外国語学習の楽しさ(個人的側面), FLEt = 外国語学習の楽しさ(教師的側面), FLEs = 外国語学習の楽しさ(社会的側面), EN\_F1 = 仲間との関わりの質, EN\_F2 = 教師との関わりの質, EN\_F3 = 学習への努力の強さ, EN\_F4 = 教材内容への取り組みの質

## 5.2 研究課題1

ここでは、研究課題を検討するために、因子に対する項目の負荷量、クロンバックのアルファ( $\alpha$ )と合成信頼性(CR), AVE, およびHTMTをそれぞれ報告する。

最初に、各因子に対する項目の因子負荷量を表2に示す。一般的に0.7以上が推奨される基準(Hair & Alamer, 2022)を14項目が下回ったものの、いずれも0.6以上であり、その他の指標も良好で、Hair et al.(2022)が示す許容範囲(0.4-0.7)に収まっていた。

■表2: 各項目の因子負荷量

| 項目    | 因子               | 因子負荷量 | t-値   | 95% CI       |
|-------|------------------|-------|-------|--------------|
| BPN1  | 自律性              | 0.761 | 23.68 | [.694, .820] |
| BPN2  |                  | 0.788 | 23.45 | [.715, .846] |
| BPN3  |                  | 0.807 | 30.50 | [.750, .853] |
| BPN4  |                  | 0.814 | 28.15 | [.752, .865] |
| BPN5  | 有能性              | 0.806 | 25.50 | [.738, .862] |
| BPN6  |                  | 0.703 | 17.70 | [.617, .772] |
| BPN7  |                  | 0.732 | 19.15 | [.649, .798] |
| BPN8  |                  | 0.731 | 20.04 | [.655, .798] |
| BPN9  |                  | 0.677 | 18.97 | [.603, .742] |
| BPN10 | 関係性              | 0.867 | 42.45 | [.823, .903] |
| BPN11 |                  | 0.716 | 17.31 | [.629, .791] |
| BPN12 |                  | 0.654 | 14.27 | [.555, .736] |
| FLE1  | 外国語学習の楽しさ(個人的側面) | 0.658 | 16.02 | [.574, .734] |
| FLE2  |                  | 0.827 | 29.77 | [.767, .877] |
| FLE3  |                  | 0.708 | 21.42 | [.639, .768] |
| FLE4  |                  | 0.698 | 19.09 | [.623, .766] |
| FLE5  |                  | 0.796 | 30.43 | [.741, .843] |
| FLE6  |                  | 0.647 | 18.19 | [.574, .713] |
| FLE8  |                  | 0.696 | 14.50 | [.595, .783] |
| FLE9  |                  | 0.804 | 25.77 | [.736, .860] |
| FLE10 | 外国語学習の楽しさ(教師的側面) | 0.819 | 26.35 | [.753, .873] |
| FLE11 |                  | 0.809 | 18.82 | [.716, .885] |
| FLE7  | 外国語学習の楽しさ(社会的側面) | 0.734 | 18.29 | [.647, .803] |
| FLE12 |                  | 0.791 | 22.97 | [.717, .851] |
| FLE13 |                  | 0.853 | 31.10 | [.798, .906] |
| FLE14 |                  | 0.726 | 16.70 | [.631, .802] |
| FLE15 |                  | 0.728 | 17.59 | [.640, .802] |
| EN1   | 仲間との関わりの質        | 0.691 | 17.42 | [.607, .761] |
| EN2   |                  | 0.775 | 24.99 | [.711, .831] |
| EN3   |                  | 0.755 | 23.65 | [.687, .812] |
| EN4   |                  | 0.718 | 17.06 | [.628, .794] |
| EN5   |                  | 0.689 | 17.51 | [.605, .761] |
| EN6   |                  | 0.791 | 24.61 | [.721, .846] |
| EN7   |                  | 0.742 | 22.14 | [.671, .803] |
| EN8   |                  | 0.689 | 18.01 | [.609, .758] |
| EN9   | 教師との関わりの質        | 0.767 | 16.99 | [.670, .844] |
| EN10  |                  | 0.647 | 11.39 | [.525, .748] |
| EN11  |                  | 0.772 | 24.49 | [.704, .827] |
| EN12  |                  | 0.825 | 29.19 | [.765, .875] |

| 項目   | 因子           | 因子負荷量 | t 値   | 95% CI       |
|------|--------------|-------|-------|--------------|
| EN13 | 教師との関わりの質    | 0.727 | 16.76 | [.630, .800] |
| EN14 |              | 0.762 | 19.34 | [.676, .830] |
| EN15 | 学習への努力の強さ    | 0.723 | 18.60 | [.642, .793] |
| EN16 |              | 0.696 | 16.89 | [.609, .771] |
| EN17 |              | 0.713 | 19.17 | [.635, .779] |
| EN18 |              | 0.664 | 15.21 | [.572, .744] |
| EN19 |              | 0.735 | 22.19 | [.665, .795] |
| EN20 |              | 0.711 | 19.09 | [.633, .778] |
| EN21 |              | 0.689 | 15.04 | [.593, .772] |
| EN22 | 教材内容への取り組みの質 | 0.696 | 15.26 | [.602, .780] |
| EN23 |              | 0.715 | 18.82 | [.634, .783] |

表3は、各因子の信頼性および収束的妥当性を判断する数値である AVE をまとめたものである。なお、詳細は後述するが、弁別的妥当性の問題に直面したため、基本的心理欲求とエンゲージメントは一つの因子としてまとめて算出した。信頼性について、すべての構成概念においてクロンバックの  $\alpha$  が基準値 0.7 を超えた。具体的には、それぞれ、エンゲージメント ( $\alpha = 0.963$ )、基本的心理欲求 ( $\alpha = 0.941$ )、外国語学習の楽しさ(個人的側面) ( $\alpha = 0.883$ )、外国語学習の楽しさ(社会的側面) ( $\alpha = 0.878$ )、外国語学習の楽しさ(教師的側面) ( $\alpha = 0.852$ ) であった。合成信頼性 (CR) においても、すべての構成概念が基準値 0.7 を上回り、 $\alpha$  とほぼ同等の値を示した。AVE については、すべての構成概念が基準値 0.5 を上回る結果となった。

■表3: 各因子の信頼性および収束妥当性

| 因子               | クロンバック $\alpha$ | 合成信頼性 | AVE   |
|------------------|-----------------|-------|-------|
| 基本的心理欲求          | 0.941           | 0.941 | 0.573 |
| 外国語学習の楽しさ(個人的側面) | 0.883           | 0.883 | 0.521 |
| 外国語学習の楽しさ(教師的側面) | 0.852           | 0.852 | 0.657 |
| 外国語学習の楽しさ(社会的側面) | 0.878           | 0.877 | 0.590 |
| エンゲージメント         | 0.963           | 0.963 | 0.529 |

最後に、弁別的妥当性について、その指標となる HTMT 値を表4に提示する。弁別的妥当性の HTMT 基準値は 0.90 未満である。分析の結果、基本的心理欲求とエンゲージメントなど多くの因子間で基準値 0.90 を上回る値が確認され、弁別的妥当性は担保されなかった。

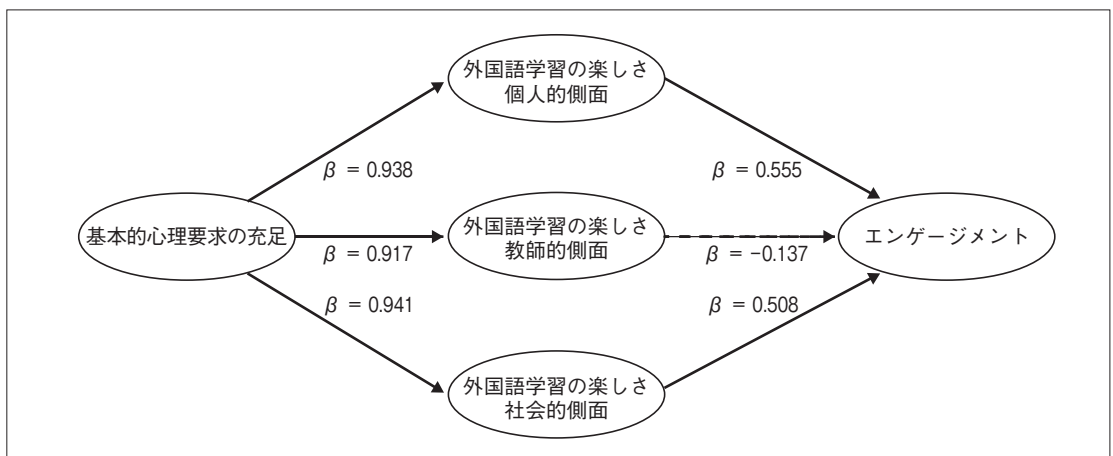
■表4: 各因子間の HTMT 値

|                  | 基本的心理欲求 | 外国語学習の<br>楽しさ<br>(個人的側面) | 外国語学習の<br>楽しさ<br>(教師的側面) | 外国語学習の<br>楽しさ<br>(社会的側面) | エンゲージメント |
|------------------|---------|--------------------------|--------------------------|--------------------------|----------|
| 基本的心理欲求          | -       |                          |                          |                          |          |
| 外国語学習の楽しさ(個人的側面) | 0.941   | -                        |                          |                          |          |
| 外国語学習の楽しさ(教師的側面) | 0.912   | 0.766                    | -                        |                          |          |
| 外国語学習の楽しさ(社会的側面) | 0.937   | 0.798                    | 0.895                    | -                        |          |
| エンゲージメント         | 0.988   | 0.956                    | 0.817                    | 0.907                    | -        |

### 5.3 研究課題2

構造モデルの検討は、パス係数、決定係数( $R^2$ )、およびサンプル外予測力を検証するためのPLS予測を用いた評価を報告する。透明性の観点から、研究課題1で言及した弁別的妥当性の問題や概念間に強い相関があり、理論的に妥当な範囲を超える数値が算出されたため、最終モデルに至るまでに複数回のモデルの調整を行ったことを報告する。最終的に、理論的妥当性と統計的適合度の両方を満たすモデルとして、基本的心理欲求の充足とエンゲージメントをそれぞれ単一の潜在変数とし、外国語学習の楽しさを三つの側面(個人的側面、教師的側面、社会的側面)に分けた現在のモデル(図2参照)を採用した。

パス係数に関して、基本的心理欲求の充足が外国語学習の楽しさの各側面に対して一貫して強い正の影響を示した。具体的には、基本的心理欲求の充足から外国語学習の楽しさ(社会的側面)への影響が最も強く( $\beta = 0.941, p < .001$ )、続いて外国語学習の楽しさ(個人的側面)への影響( $\beta = 0.938, p < .001$ )、外国語学習の楽しさ(教師的側面)への影響( $\beta = 0.917, p < .001$ )の順で観察された。エンゲージメントに対する各外国語学習の楽しさ要因の影響については、異なる結果が見られた。外国語学習の楽しさ(個人的側面)はエンゲージメントに対して中程度の正の影響( $\beta = 0.555, p < .001$ )を示し、外国語学習の楽しさ(社会的側面)も同程度の正の影響( $\beta = 0.508, p < .001$ )を与えていた。対照的に、外国語学習の楽しさ(教師的側面)はエンゲージメントに対して弱い負の影響( $\beta = -0.137, p < .05$ )を示した。



■図2: 構造モデルの結果

決定係数は、エンゲージメントの分散の96.5%を説明し( $R^2 = .965$ )、また、各外国語学習の楽しさの各要素についても高い説明力が確認された。外国語学習の楽しさ(個人的側面)( $R^2 = .880$ )、外国語学習の楽しさ(教師的側面)( $R^2 = .841$ )、外国語学習の楽しさ(社会的側面)( $R^2 = .885$ )。これらの結果は、基本的心理欲求の充足が外国語学習における楽しさを包括的に説明することを示唆している。予測妥当性について、実際の観測値を反映したPLS-SEMと単純線形回帰モデル(LM)の二乗平均平方根誤差(RMSE)を比較した結果、大部分の指標において、PLS-SEMの値はLMの値よりも低い予測誤差を示した(例: FLE1 [PLS = 0.853 < LM = 0.858])。これは本モデルが中程度の予測力を持つことを示している(Hair & Alamer, 2022)。媒介分析について、基本的心理欲求からエンゲージメントへの総効果が非常に大きいことが確認された( $\beta = 0.965$ )。間接効果の詳細を確認すると、最も強い間接効果は基本的心理欲求から外国語学習の楽しさ(個人的側面)を経由してエンゲージメントに至るもので、その効果は $\beta = 0.614$ であった。次に、基本的心理欲求から外国語学習の楽しさ(社会的側面)を経由する経路が中程度の強さを示し、間接効果は $\beta = 0.476$ であった。一方、基本的心理欲求から外国語学習の楽しさ(教師的側面)を経由する経路は負の間接効果( $\beta = -0.125$ )を示した。

## 6 考察

本研究では、児童を対象に外国語学習の楽しさを媒介として、基本的心理欲求の充足がエンゲージメントに与える影響を検証した。上述の結果に基づき、以下では研究課題を中心に考察を行う。

### 6.1 研究課題1:「基本的心理欲求・外国語学習の楽しさ・エンゲージメントの測定モデルは、どの程度、信頼性および妥当性が確保されているか。」

Hair and Alamer (2022) の部分最小二乗構造方程式モデリングのガイドラインに従って測定モデルを検証した結果、信頼性指標(クロンバック  $\alpha > .70$ , 合成信頼性  $> .70$ )および収束的妥当性指標(AVE  $> .50$ )については、大部分の構成概念で基準値を満たしていた。しかし、HTMTを指標とする弁別的妥当性においては、基準値の0.9を大幅に上回る値が多く観察された。弁別的妥当性とは、Hair et al. (2023)によると“the extent to which a construct is truly distinct from other constructs by empirical standards. Thus, establishing discriminant validity implies that a construct is unique and captures phenomena not represented by other constructs in the model” (p. 120)と定義されている。つまり、本研究のHTMT値は、基本的心理欲求、外国語学習の楽しさ、そして、エンゲージメントの各構成概念が実質的に同一の要素を測定している可能性が示唆される。この弁別的妥当性が満たされなかった点について、以下の三つの観点から考察する。

一つ目に、基本的心理欲求、外国語学習の楽しさ、エンゲージメントの尺度開発過程における手続き的要因が考えられる。本研究で用いた構成概念は理論的には異なるものの、先行研究から概念的な近さは想定されていた。その近い概念が、成人用尺度を児童用に改訂する過程で項目内容がさらに類似してしまった可能性が考えられる。つまり、児童の発達段階に合わせて語彙を調整したり、文構造を簡単にするという改訂作業により、本来異なる理論的概念を表していた項目が、表面的に類似した表現に収束してしまった可能性が考えられる。その結果、それぞれの項目をランダムに配列して組み込んだ50項目からなる質問紙において、理論的に区別されるべき三つの構成概念間の境界が曖昧になったと考えられる。

二つ目に、この研究では小学生を対象としており、アンケート調査ができる年齢の中でも下限の年齢の児童を対象に行った研究である。基本的心理欲求、外国語学習の楽しさ、エンゲージメントといった複雑で抽象的な概念は、まだ認知発達の途中にある小学生にとって、これらの細かい概念の違いを理解し、それぞれを区別して回答することは認知的に困難であった可能性が高い(Butler, 2018)。すなわち、子どもたちはこれらの概念を区別せず、一体的な学習体験として捉えていた可能性がある。

三つ目に、児童の回答行動が考えられる。データクリーニングの結果、38名(全体の約11%)が全項目に同一の回答をしていたことが確認された。これは、50項目という項目数が児童にとって過大な認知的負荷となり、回答の質を著しく低下させた可能性を示している。結果として、質問内容を適切に判断することなく、機械的に回答した可能性が考えられる。

以上の考察から、本研究において、HTMT値が示す弁別的妥当性の問題は、質問紙改訂の際に生じる手続き的問題、子どもの認知的発達段階、複雑なモデルの回答への負荷などが複合的に影響した結果であると考えられる。

### 6.2 研究課題2:「基本的心理欲求の充足が外国語学習の楽しさを媒介としてエンゲージメントに与える影響はどの程度か。」

本研究の結果を解釈する際には、研究課題1で示された弁別的妥当性の課題により、深い議論を展開することはできず、限定的な解釈にとどまらざるを得ない。この前提の上で、本研究の主要な発見について検討する。まず、基本的心理欲求の充足が外国語学習の楽しさのすべての側面に対して強い正の影響を示した点は、外国語学習の楽しさが基本的心理欲求の充足を前提とするという理論的背景と一致してい

た。すなわち、基本的心理欲求が満たされることで、学習者は外国語学習を楽しいと感じやすくなると考えられる。一方で、外国語学習の楽しさがエンゲージメントに及ぼす影響は、楽しさの側面によって異なることが明らかになった。具体的には、外国語学習の楽しさの個人的側面および社会的側面はエンゲージメントを促進する一方で、教師の側面は弱いながらも負の影響を示した。もっとも、前述のように弁別的妥当性の問題があるため、これらの結果の解釈には慎重さが求められる。

## 7 結論

本研究では、小学校5・6年生を対象に、基本的心理欲求、外国語学習の楽しさ、学習エンゲージメントの関係性を部分最小二乗構造方程式モデリングにより検証した。その結果、これらの概念間には理論的に予想される関係性が確認された一方で、重要な方法論的課題も明らかとなった。本研究で示された弁別的妥当性の問題は、単なる研究上の限界にとどまらず、児童を対象とした質問紙調査法に対する重要な示唆を提供するものである。HTMT値に基づく高い概念間相関は、近接する複数の心理的概念を同時に測定することの根本的な難しさを示しており、児童を対象とした質問紙調査の適用範囲を再考する必要性を浮き彫りにした。一方で、外国語学習の楽しさのような単一概念の測定においては、質問紙調査が十分に機能し、その妥当性と信頼性が確保されることも確認された。したがって、基本心理的欲求、外国語学習の楽しさ、エンゲージメントといった理論的に近接する複数概念を同一調査で扱う際には、弁別的妥当性の確保が大きな課題となることが実証的に示されたといえる。最後に、このような方法論的自省は研究の質を高め、実践への貢献を可能にする。本研究の成果と課題が、今後の児童を対象とした研究の発展に寄与することを期待したい。

### 謝辞

はじめに、公益財団法人 日本英語検定協会の皆様ならびに選考委員会の皆様に、本研究の実施の機会を賜りましたことを心より感謝申し上げます。特に、研究助言者である寺内 一先生には、貴重なご助言を賜りましたことに深く御礼申し上げます。また、本研究にご協力くださいました小学校の先生方、そして児童の皆様にも厚く感謝申し上げます。なお、本報告書に関連して開示すべき利益相反事項はございません。

### 倫理的配慮に関する指針

本研究では、調査実施に先立ち、筆者が在籍する関西大学外国語教育学研究科の倫理審査を受け、その承認を得た上で、調査を依頼した各学校の校長より承諾を得た。使用した質問紙は無記名形式とし、得られたデータは研究目的以外には使用せず、個人が特定される情報は一切収集しないこととした。また、収集したデータについては、第三者に知られることのないよう厳重に管理した。質問項目については、研究者側で児童にとって負担のない内容であることを事前に確認し、さらに予備調査を依頼した管理職、英語専科教員、および担任教員に内容を精査してもらった。特に、児童の心理的負担を軽減するため、否定的表現をすべて肯定的表現に変更するなど、尺度の妥当性および信頼性を担保しつつ、適切な倫理的配慮を行った。

## 引用文献

- Agawa, T., & Takeuchi, O. (2016). A new questionnaire to assess Japanese EFL learners' motivation: Development and validation. *Annual Review of English Language Education in Japan*, 27, 1-16. [https://doi.org/10.20581/arele.27.0\\_1](https://doi.org/10.20581/arele.27.0_1)
- Alamer, A., Robat, E. S., Shirvan, M. E., & Ryan, R. (2025). Self-determination theory and language learning: A multilevel meta-analysis. *Educational Psychology Review*, 37(2), 59. <https://doi.org/10.1007/s10648-025-10038-y>
- Al-Hoorie, A. H., Oga-Baldwin, W. L. Q., Hiver, P., & Vitta, J. P. (2025). Self-determination mini-theories in second language learning: A systematic review of three decades of research. *Language Teaching Research*, 29(4), 1603-1638. <https://doi.org/10.1177/13621688221102686>
- Alrabai, F., & Algazzaz, W. (2025). Testing experimental-based models on the influence of teacher emotional support on students' basic psychological needs, emotions, and emotional engagement. *Applied Linguistics*. <https://doi.org/10.1093/applin/amaf036>
- Andringa, S., & Godfroid, A. (2020). Sampling bias and the problem of generalizability in applied linguistics. *Annual Review of Applied Linguistics*, 40, 134-142. <https://doi.org/10.1017/s0267190520000033>
- Boo, Z., Dörnyei, Z., & Ryan, S. (2015). L2 motivation research 2005-2014: Understanding a publication surge and a changing landscape. *System*, 55, 145-157. <https://doi.org/10.1016/j.system.2015.10.006>
- Botes, E., Dewaele, J., & Greiff, S. (2021). The development and validation of the short form of the foreign language enjoyment scale. *The Modern Language Journal*, 105(4), 858-876. <https://doi.org/10.1111/modl.12741>
- Botes, E., Dewaele, J.-M., & Greiff, S. (2022). Taking stock: A meta-analysis of the effects of foreign language enjoyment. *Studies in Second Language Learning and Teaching*, 12(2), 205-232. <https://doi.org/10.14746/sslit.2022.12.2.3>
- Butler, Y. G. (2018). The role of context in young learners' processes for responding to self - assessment items. *The Modern Language Journal*, 102(1), 242-261. <https://doi.org/10.1111/modl.12459>
- Csikszentmihalyi, M. (2008). *Flow : the psychology of optimal experience*. Harper Perennial.
- Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237-274. <https://doi.org/10.14746/sslit.2014.4.2.5>
- Dewaele, J.-M., & MacIntyre, P. D. (2016). Foreign language enjoyment and foreign language classroom anxiety. The right and left feet of FL learning? In P. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215-236). Multilingual Matters. <https://doi.org/10.21832/9781783095360-010>
- Eerdemutu, L., Dewaele, J.-M., & Wang, J. (2024). Developing a short language classroom engagement scale (LCES) and linking it with needs satisfaction and achievement. *System*, 120, 103189. <https://doi.org/10.1016/j.system.2023.103189>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, State of the evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>
- Hair, J., & Alamer, A. (2022). Partial Least Squares Structural Equation Modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics*, 1(3), 100027. <https://doi.org/10.1016/j.rmal.2022.100027>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. 3rd Edition. Sage.
- Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, 28(1), 201-230. <https://doi.org/10.1177/13621688211001289>
- Hosseini, H. M., Fathi, J., Derakhshesh, A., & Mehraein, S. (2022). A model of classroom social climate, foreign language enjoyment, and student engagement among English as a foreign language learners. *Frontiers in Psychology*, 13, 933842. <https://doi.org/10.3389/fpsyg.2022.933842>
- Jin, Y. (2024). Motivating students to actively engage in EFL classrooms: Exploring the role of L2 grit and foreign language enjoyment. *Learning and Motivation*, 85. <https://doi.org/10.1016/j.lmot.2024.101960>
- Khajavy, G. H. (2021). Modeling the relations between foreign language engagement, emotions, grit and reading achievement. In P. Hiver, A. H. Al-Hoorie, & S. Mercer (Eds.), *Student Engagement in the Language Classroom* (pp. 241-259). Multilingual Matters.
- Li, C., Jiang, G., & Dewaele, J.-M. (2018). Understanding Chinese high school students' foreign language enjoyment: Validation of the Chinese version of the foreign language enjoyment scale. *System*, 76, 183-196. <https://doi.org/10.1016/j.system.2018.06.004>
- MacIntyre, P. D., & Vincze, L. (2017). Positive and negative emotions underlie motivation for L2 learning. *Studies in Second Language Learning and Teaching*, 7(1), 61-88. <https://doi.org/10.14746/sslit.2017.7.1.4>
- MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching*, 2(2), 193-213. <https://doi.org/10.14746/sslit.2012.2.2.4>
- Martin, A. J., Ginns, P., & Papworth, B. (2017). Motivation and engagement: Same or different? Does it matter? *Learning and Individual Differences*, 55, 150-162. <https://doi.org/10.1016/j.lindif.2017.03.013>
- 文部科学省 (2017) 小学校学習指導要領(平成29年告示)解説 外国語活動・外国語編
- Nakanishi, Y., & Takeuchi, O. (2024). Validating the foreign language enjoyment scale for young learners: An exploratory structural equation modeling approach. *Research Methods in Applied Linguistics*, 3(3), 100167. <https://doi.org/10.1016/>

## 引用文献

- j.rmal.2024.100167
- Noels, K. A., Clément, R., & Pelletier, L. G. (1999). Perceptions of teachers' communicative style and students' intrinsic and extrinsic motivation. *The Modern Language Journal*, 83(1), 23-34. <https://doi.org/10.1111/0026-7902.00003>
- Sang, Y., & Hiver, P. (2021). Engagement and companion constructs in language learning: Conceptualizing learners' involvement in the L2 classroom. In P. Hiver, A. H. Al-Hoorie, & S. Mercer (Eds.), *Student Engagement in the Language Classroom* (pp. 17-37). Multilingual Matters.
- Shirvan, M. E., Taherian, T., & Yazdanmehr, E. (2024). Foreign language enjoyment: a longitudinal confirmatory factor analysis-curve of factors model. *Journal of Multilingual and Multicultural Development*, 45(2), 224-242. <https://doi.org/10.1080/01434632.2021.1874392>
- Sudina, E. (2021). Study and scale quality in second language survey research, 2009-2019: The case of anxiety and motivation. *Language Learning*, 71(4), 1149-1193. <https://doi.org/10.1111/lang.12468>
- Sugita-McEown, M., & Oga-Baldwin, W. L. Q. (2019). Self-determination for all language learners: New applications for formal language education. *System*, 86, 102124. <https://doi.org/10.1016/j.system.2019.102124>
- Teimouri, Y., Goetze, J., & Plonsky, L. (2023). Second language anxiety and achievement. *Studies in Second Language Acquisition*, 41(2), 363-387. <https://doi.org/10.1017/s0272263118000311>
- Teravainen-Goff, A. (2022). Why motivated learners might not engage in language learning: An exploratory interview study of language learners and teachers. *Language Teaching Research*, 29(6), 2500-2520. <https://doi.org/10.1177/13621688221135399>
- Teravainen-Goff, A. (2023). Intensity and perceived quality of L2 engagement: Developing a questionnaire and exploring engagement of secondary school language learners in England. *System*, 112, 102955. <https://doi.org/10.1016/j.system.2022.102955>
- Tsang, A., & Dewaele, J. M. (2023). The relationships between young FL learners' classroom emotions (anxiety, boredom, & enjoyment), engagement, and FL proficiency. *Applied Linguistics Review*, 5(5), 2015-2034. <https://doi.org/10.1515/applrev-2022-0077>
- Vo, H. (2024). Towards clarification of the second language learner engagement construct: Taking stock of its conceptual overlap and hierarchical structure. *Applied Linguistics*, 45(6), 1030-1049. <https://doi.org/10.1093/applin/amae012>
- Wang, N. (2024). How does basic psychological needs satisfaction contribute to EFL learners' achievement and positive emotions? The mediating role of L2 self-concept. *System*, 123, 103340. <https://doi.org/10.1016/j.system.2024.103340>
- Zhang, X. (2019). Foreign language anxiety and foreign language performance: A meta-analysis. *The Modern Language Journal*, 103(4), 763-781. <https://doi.org/10.1111/modl.12590>
- Zhou, S. A., Hiver, P., & Al-Hoorie, A. H. (2021). Measuring L2 engagement: A Review of issues and applications. In P. Hiver, A. H. Al-Hoorie, & S. Mercer (Eds.), *Student Engagement in the Language Classroom*. Multilingual Matters.

## 資料1: 質問紙項目一覧

|    | 質問項目                                   | 分類    |
|----|----------------------------------------|-------|
| 1  | 私は英語の授業では、自由に新しいアイデアを考えることができる。        | FLE1  |
| 2  | 私は英語の授業中、楽しく英語を学んでいる。                  | FLE2  |
| 3  | 私は英語の授業で、自分が言いたいことを英語で伝えられるようになってきた。   | FLE3  |
| 4  | 英語の授業では、自分のがんばりがみんなの役に立っていると思う。        | FLE4  |
| 5  | 私は英語の授業では、興味深いこと(もっと知りたいと思えること)を学んでいる。 | FLE5  |
| 6  | 私は英語の授業では、自分ができたと思えることが多い。             | FLE6  |
| 7  | 英語の授業は、明るくポジティブ(前向き)なふんいきである。          | FLE7  |
| 8  | 私は英語の授業では、まちがうことも、次につながる学びだと思う。        | FLE8  |
| 9  | 英語の先生はいつもはげまそうとしてくれる。                  | FLE9  |
| 10 | 英語の先生は、フレンドリー(親しみやすい)に接してくれる。          | FLE10 |
| 11 | 英語の先生は、サポートしようと(助けようと)してくれる。           | FLE11 |
| 12 | 英語の授業は、良いふんいきだ。                        | FLE12 |
| 13 | 私たちの英語のクラスには、一体感(まとまり)がある。             | FLE13 |

## (資料1 続き)

|    | 質問項目                                         | 分類    |
|----|----------------------------------------------|-------|
| 14 | 私たちのクラスには、みんなが笑顔になれるしゅんかんがある。                | FLE14 |
| 15 | 私たちは、英語の授業中によく笑いあう。                          | FLE15 |
| 16 | 英語の授業では、先生は私たちの意見や発言をよく聞いてくれる。               | BPN1  |
| 17 | 英語の授業で、先生は授業中に行う活動やその内容の大切さを説明してくれる。         | BPN2  |
| 18 | 英語の先生は、私たちの気持ちをわかってくれていると思う。                 | BPN3  |
| 19 | 英語の先生は、英語の学習がうまくいくようにはげましてくれる。               | BPN4  |
| 20 | 英語の授業で、先生は私たちの意見を大切にしてくれていると思う。              | BPN5  |
| 21 | 英語の先生は、質問しやすいと思う。                            | BPN6  |
| 22 | 英語の授業では、自分の努力してきたことがうまくいったと感ずることがある。         | BPN7  |
| 23 | 英語の授業では、「できた」という気持ちになることがある。                 | BPN8  |
| 24 | 英語の授業では、自分の頑張りにも満足している。                      | BPN9  |
| 25 | 英語の授業では、楽しくて心地よい雰囲気(ふんいき)があると思う。             | BPN10 |
| 26 | 英語の授業では、クラスの仲間と協力してやっていると思う。                 | BPN11 |
| 27 | 英語の授業では、グループ活動やペアワークで、協力がしやすいと思う。            | BPN12 |
| 28 | 私はグループ活動(班活動)の時、メンバーと一緒にいて居心地(いごち)が良いと感じる。   | EN1   |
| 29 | 私は普段から、クラスメートと一緒に活動することで、多くのことを学べていると感じている。  | EN2   |
| 30 | 私は普段から、英語の授業で行う活動をクラスメートと一緒にやり切ることを楽しみにしている。 | EN3   |
| 31 | 私は普段から、グループ活動(班活動)は英語を学習するのに良い方法だと感じている。     | EN4   |
| 32 | 私は普段から、クラスメートと一緒に協力して学んでいる。                  | EN5   |
| 33 | 私は普段から、クラスの雰囲気(ふんいき)がより良い学びにつながると感じている。      | EN6   |
| 34 | 私は普段から、グループ活動(班活動)の時には一生懸命に取り組んでいる。          | EN7   |
| 35 | 私は普段から、授業中、クラス全体での話し合いに参加している。               | EN8   |
| 36 | 英語の先生は普段から、私の英語学習がうまくいくように助けてくれる。            | EN9   |
| 37 | 英語の先生は普段から、英語学習に役立つヒントや手助けを与えてくれる。           | EN10  |
| 38 | 私は普段から、英語の先生が、私のやりたい事を聞き入れてくれると感じている。        | EN11  |
| 39 | 私は普段から、英語の先生の授業のやり方に満足している。                  | EN12  |
| 40 | 普段から、英語の先生と良い関係である。                          | EN13  |
| 41 | 私は普段から、英語の先生からたくさんを学んでいると感じている。              | EN14  |
| 42 | 私は普段から、授業中に行われる活動に集中している。                    | EN15  |
| 43 | 私は普段から、授業中の学習や活動には全力で取り組む。                   | EN16  |
| 44 | 私は、授業中にやるのが難しいと感じても、あきらめない。                  | EN17  |
| 45 | 私は普段から、英語の先生の話をしっかり聞いている。                    | EN18  |
| 46 | 私は普段から、授業の内容をわかろうと努力している。                    | EN19  |
| 47 | 私は普段から、自分にとって良い学び方を考えている。                    | EN20  |
| 48 | 私は普段から、授業で使われる教材等は、役立つと感じている。                | EN21  |
| 49 | 授業で扱われる教材等(教科書など)は、英語を学ぶ上で役に立つ。              | EN22  |
| 50 | 授業の内容は、私の学びにつながると感じる。                        | EN23  |

## 第37回 研究助成

## A 研究部門・報告V・英語能力テストに関する研究

# 多肢選択式読解問題における本文の語を含む割合を調整した正答選択肢の調査

## ー 生成 AI を用いた選択肢の作成方法の検討 ー

研究者: 若松 千智 千葉県／千葉大学大学院 在籍

《研究助言者: 齊田 智里》

### 概要

本研究の目的は、多肢選択式読解問題の正答選択肢における「本文の語を含む割合」を調整した場合、項目正答率・自信度・弁別力のそれぞれがどのように変化するのか検証することである。また、テスト作成者である教員の負担軽減のために、生成 AI を用いた効率的な選択肢の作成方法を検討した。英検2級の読解問題(オリジナル)と、正答選択肢のみを生成 AI で再作成した問題(パラフレーズ)の二種類のテストを国立大学・大学院に所属する65名に実施した。その結果、本文の語を含む割合を低くした正答選択肢を用いた場合、自信度・弁別力ともに有意差は見られなかった。一方で、本文中の語を多く含む正答選択肢を用いた場合と比較して、正答選択肢に含まれる本文中の語の割合が低い場合には、項目正答率が有意に低下することが確認された。また、いずれのテストにおいても高い信頼性が示された。以上より、生成 AI を用いたパラフレーズは項目の信頼性や弁別力を損なわずに活用可能であり、効率的な問題作成に有用な可能性が示唆された。

## 1 はじめに

### 1.1 学習指導要領

グローバル化が進んでいく中、小学校で外国語が教科化されるなど、英語の重要性は高まっている。平成29年、平成30年告示の学習指導要領解説・外国語編では、生徒が英語に触れる機会を充実するとともに、授業を実際のコミュニケーションの場面にするため、授業は英語で行うことを基本とすると定められており、リーディングにおいても英語を英語のまま理解することは、読みの流暢性の観点からも重要である。卯城(2011)では、英文には必ずそれを書くための目的があり、英文を理解することは読み手が書き手のメッセージを受け取ることでありと言及されており、一文一文を理解することができても、必ずしも英文全体のメッセージをつかむことはできないと言われている。英文を理解するためには、語彙や文法の知識だけではなく、背景知識を活性化させたり能動的に読んだりすることで文章全体のメッセージを捉えることが求められる。授業を実際のコミュニケーションの場面にするため、英語の授業は英語で行うことを基本としていることから、英語で内容を理解することが求められている。そのため、実際の定期テストでも英語を用いての内容理解を問うべきだが、記述式では中学生にとって難易度が高くなってしまいう懸念がある。

指導と評価のずれとして「知識及び技能」と「思考力、判断力、表現力等」の境界線の問題が今井(2023)で挙げられている。同文献では、「知識及び技能」と「思考力、判断力、表現力等」は区別が難しいことが指摘されているものの、目的・場面・状況をもとに言語活動を設定し、適切に言語使用することが「思考力、

判断力、表現力等」であり、正確さが「知識及び技能」と定義されている。評価方法にパフォーマンステストを取り入れることにより、紙のテストの限界を補うことは重要であるものの、4技能のうち読解能力を測るためにはやはり紙のテストの充実が求められる。しかし、紙の定期テストでは、観点別評価をするために「知識及び技能」を測る問題と「思考力、判断力、表現力等」を測る問題を分ける必要があり、中でも読解問題における「思考力、判断力、表現力等」を測る問題の作成が課題である。

## 1.2 生成AIについて

ChatGPTが2022年にリリースされて以降、教育の在り方にも影響が出ている。初等中等教育における生成AIの利活用に関するガイドラインver.2.0(文部科学省, 2024)によると、児童生徒が学習に生成AIを利活用する際だけでなく、教員が生成AIを校務で利活用することに関するガイドラインが示されている。まず、教育は教員と児童生徒の人格的な触れ合いを通じて行われるものであり、生成AIが社会インフラの一部となる時代においてより重要になる。しかし、教員が授業準備や文書の叩き台作成を含む校務において生成AIを利活用することで校務の効率化につなげることができると示されている。さらに、教員が生成AIの利便性や懸念点を知ることが、児童生徒の学びをより高度化する観点からも重要であると言われている。教員自身が生成AIの仕組みや活用方法について学ぶことで業務の効率化につなげることができるが、不正確な情報を出力してしまうハルシネーションのおそれもあるため、最終的には教員が生成AIからの出力を正しく判断する必要がある。

生成AIには、OpenAI社のChatGPT、Google社のGemini、Microsoft社のCopilotをはじめとして、数多く開発されている。日本リサーチセンターの調査によると、生成AIの利用割合は2024年6月から2025年6月までに14.7パーセンテージポイント上昇し、2025年6月には30.3%( $n = 1,394$ )を占めている。生成AIを使用している人が利用している生成AIの種類は、上記3種類の中でChatGPTの利用率は20.9%と最も高い。ChatGPTの仕組みについて、大規模コーパスを使用し、事前に言語使用の深層学習を行っており、次に来る確率の高い語やフレーズを生成していると水本(2024)で説明がされている。このことから、生成AIは言語処理が得意であり、英語教育においても親和性が高いと言える。日本の大学入学共通テストにおいて生成AIを活用した研究としてWatanabe and Saida(2024)が挙げられる。Watanabe and Saida(2024)では、大学入学共通テストの英語リーディングの問題を生成AIがどの程度得点できるかを調査した。その結果、ChatGPTの平均得点は受験者平均を大きく上回ることが明らかになった。さらに、ChatGPTは設問の不明瞭さを発見するなど項目の改善に役立つ可能性が示唆されたことから、英語教員の補助ツールとしてテスト検証や教材作成に有効であると結論づけている。また、生成AIを使用したテスト作成について検討したShin(2023)では、生成AIを用いて多肢選択式テストを作成し、実際に使用されている韓国の大規模試験(CSAT)との比較を50名の英語教員が行った。研究協力者である英語教員は、ChatGPTを使用して作成したテストと実際に使用されている韓国の大規模試験において、文の流れや表現の自然さを同程度と評価したことが示されている。これらの結果は、生成AIを使用したテスト作成の可能性を示唆している。

## 1.3 テスト作成について

テストを検討する際に意識する必要がある妥当性・信頼性・波及効果・実用性の4観点について、根岸(2017)では以下のように定義している。

- 妥当性……テストが測ろうとしている知識や能力を測っているか
- 信頼性……何度測っても同じような結果が出るか
- 波及効果……テストが学習や指導に及ぼす影響はどのようなものか
- 実用性……テストの作成・実施・採点・解釈が容易かどうか

これらの4観点が十分に反映されたテスト作成が重要だが、近年の教育現場では教員不足や教員の多忙化が著しく、妥当性、信頼性、波及効果の3つが保たれた上での「実用性」の向上が求められる。定期テストの読解問題の課題として、根岸(2017)では、定期テストでは教科書本文を読解問題として出題しているが、すでに読んだことのある文章はテストで改めて読まなくても内容がわかってしまうことが挙げられていることから、教科書や授業では扱っていない初見の文章を使用する必要性があると言える。

また、読解力を測る手段としてクローズテスト、多肢選択式テスト、組み合わせ法、空所補充、真偽テスト、要約テストなどがあるが、その中の一つである多肢選択式テストは、作成に時間や手間がかかるものの採点が容易であるため、実用英語技能検定、IELTS、TOEFL、TOEICなどの多くの大規模テストで使用されている。一方、多肢選択式テストのデメリットとして、当て推量の可能性があることが挙げられている。多肢選択式の読解問題と記述式(Open-Ended)の読解問題において、項目の種類による難易度への影響を調査したUshiro et al.(2008)では、多肢選択式テストは記述式テストと比べて、点数が予測しにくいことが明らかになっている。このことから、記述式では理解できていないものを産出することができない反面、選択式のテストは分からなくても当たってしまったり、なんとなく当ててしまったりするおそれがあると言える。そのような「当て推量」を防ぐために有効な手段として、もっともらしい錯乱肢を作成することが挙げられており、テスト受験者の能力によって思わず選んでしまう、引きつけられやすさを指す「魅力度」が変化する選択肢の検討が求められる(若林・根岸, 1993)。

多肢選択式読解問題における質問タイプを分析した清水(2008)では、質問タイプを大きく上位レベル処理と下位レベル処理の2つに分けている。上位レベル処理の質問タイプとして、推論質問、テーマ質問、文章構造質問が挙げられており、これらは下位レベル処理の質問タイプであるパラフレーズ質問よりも処理が難しいことが示されている。清水(2008)でのパラフレーズ質問は、文章内で記されている文章を言い換えたものであり、先述した今井(2023)で言われている「知識及び技能」の定義である「言語使用の正確さ」を測る項目だと判断できる。「言語使用の正確さ」に加えて、より上位レベルの処理を要する問題を取り入れることにより、「思考力、判断力、表現力等」の定義として示されていた「言語使用の適切さ」を測ることができらるだろう。Hasegawa(2017)では、正誤問題を、テキストに明示されているか、もしくは暗示されているかについて、1:推論が必要か、2:2文以上を参照する必要があるかの2つの基準で分類している。テキストに暗示されている内容に関する質問は、清水(2008)の上位レベル処理の質問タイプであり、「思考力、判断力、表現力等」を測る質問であると言える。

## 1.4 選択肢について

多肢選択式テストでの選択肢は、正答の選択肢と誤りの選択肢である錯乱肢の2種類で構成される。Ushiro et al.(2007)では、TOEFLのテストを使用し、大学生114名を対象に解答者が錯乱肢を選びやすくなる要素を検討しており、錯乱肢のもっともらしさ(plausibility)について、語数や構文などの定量的なデータで判断することは難しく、問題タイプによって異なる要素が確認されている。また、どのような錯乱肢が受験者の能力を正確に識別できるかについて調査した研究として、要約タイプの問題を使用して日本人学習者が要約の過程で起こしがちな典型的なエラーである3種類の錯乱肢を比較したTerao and Ishii(2020)が挙げられる。パッセージの要点を捉える多肢選択式の問題において、1:重要な情報が欠けていたり、不必要な情報が含まれている deletion, 2:本文中の例をそのまま抜き出したり、一般化しすぎた表現を使用した generalization, 3:段落全体ではなく、一部の内容しか要約をしていなかったり、書き手の視点とは異なる記述をしている integration, の3種類の錯乱肢を使用した。その結果、重要な情報が欠けている錯乱肢(1:deletion)と詳細な例をそのまま抜き出した要約(3:integration)はテスト受験者の能力を正確に識別することが明らかになっていた。

また、清水(2008)により分類された質問タイプの一つであるパラフレーズ質問に焦点を当てた研究(政所, 2018)では、多肢選択式読解問題のパラフレーズ質問をパラフレーズ・レベルを用いて、錯

乱肢だけでなく全ての選択肢において分析を行っている。TOEIC, TOEFL, 英検2級の3つのテストの読解問題を使用し、それぞれのテストの選択肢が本文の語をどの程度使用しているかを調査した上でテストを実施した。テスト結果から、テスト得点群によって正答・誤答選択肢の組み合わせによる選択率と魅力度への影響が異なること、特にそのテストで測る能力が低い人たち(低位群)は正答選択肢が本文の語をあまり含まない場合に、より本文の語を含む錯乱肢を選ぶ傾向があることが明らかにされている。

これらの研究から、思わず選んでしまうもっともらしい錯乱肢の要素は問題タイプに依存するものの、能力によって本文の語を含んでいる割合が選ばれやすさの要因となる可能性がある。そこで、本研究では表層的な要素である本文の語を含んでいる割合を調整して調査を行うこととした。

## 1.5 本研究

本研究において使用する言葉の定義について、項目正答率は正解を1、不正解を0として算出したテストの得点率を指し、自信度は、低中高の3段階で、低を1、中を2、高を3として算出した数字を指す。今回扱うテストで測る能力をどの程度識別することができるかを示す弁別力においては、テスト全体の得点とその項目の得点の相関を表す点双列相関係数の値を弁別力の指標として扱う。

本研究では、正答選択肢における本文の内容語を言い換えた割合を調整した場合、項目正答率、テスト受験者の正答選択肢に対する自信度、弁別力のそれぞれがどのように変化するかを検証する。また、テスト作成者である教員の負担軽減のために、多肢選択式テストの課題である「作成に時間がかかること」を克服するツールとして生成AIを用い、効率的な選択肢の作成方法を検討することを目的とし、テストの充実の一助となるよう、研究課題を以下に3つ設定した。

本文の語を使用する割合がより低い正答選択肢を含む項目は、本文の語を使用する割合がより高い正答選択肢を含む項目と比べて、

**研究課題 1** 項目正答率がどの程度変化するのか明らかにすること。

**研究課題 2** 受験者の項目に対する自信度がどの程度変化するのか明らかにすること。

**研究課題 3** 受験者の能力を識別する弁別力はどの程度変化するのか明らかにすること。

これらの研究課題に対して、表面的な難易度が上がるため、研究課題1における項目正答率と研究課題2における自信度は下がり、研究課題3における弁別力は上がると仮説を立てた。

## 2 方法

### 2.1 マテリアル

テストは実用英語技能検定(以下、英検)2級2023年度第3回実施分(大問3:A, B, C), 2024年度第1回実施分(大問3:A, B), 2024年度第2回実施分(大問3:A, B)の計3回分を使用した。英検は2024年度から出題形式が変更となり、ライティング問題が1題追加された影響でリーディングの問題数が減っているため、2023年度実施分と2024年度実施分の読解問題の数が異なっている。それぞれ大問3の文章を読んだ上でいくつかの質問に答える形式の読解問題のみを抽出し、7つの文章題で小問は合計で28項目使用した。文章

題7つは通し番号をA～Gとし、テストは2種類使用した。1つ目は英検の問題をそのまま使用したもの(オリジナル)、2つ目は英検の文章、問題文、錯乱肢はそのまま使用し、正答の選択肢のみパラフレーズしたもの(パラフレーズ)である。それぞれに対して、解答の自信度を調べるために、一つ一つの設問に対して3段階で自信度を尋ねた。

## 2.2 知識及び技能を測る項目と思考力、判断力、表現力等を測る項目の分類

知識及び技能を測る項目か思考力、判断力、表現力等を測る項目かという分類の際には、英語教育を専門とする大学院生が参加し、2名で分類を行った。分類基準は、Hasegawa(2017)、清水(2005)、国立教育政策研究所 教育課程研究センター(2020)を参考にし、(1)代名詞が表すものも含めて正答にたどり着くために2文以上参照する必要がある、(2)推論が必要である、(3)段落や文章全体の主題(テーマ)を問う質問である、(4)必要な情報を読み取り、書き手の意図、概要、要点を捉える問題である、の4つに定めた。上記の4つの基準のうち、どれか一つ以上該当する場合に思考力、判断力、表現力等を測る項目とした。全ての項目を2名で分類したところ、一致率は75%だった。2名で一致しなかった項目は話し合うことにより全ての項目を分類したところ、知識及び技能を測る項目は16項目、思考力、判断力、表現力等を測る項目は12項目だった(表1参照)。

■表1: 知識及び技能と思考力、判断力、表現力等の問題分類

| 知識及び技能 | 思考力、判断力、表現力等 |
|--------|--------------|
| 16項目   | 12項目         |

## 2.3 選択肢における本文の内容語を使用した割合の調査

政所(2018)を参照し、本文に使用されている内容語が選択肢に占める割合を事前に調査した。内容語には名詞・副詞・形容詞・一般動詞を含め、機能語にはbe動詞・助詞・助動詞・接続詞・冠詞・前置詞・代名詞・疑問詞を含めて分類を行った。一つの単語に複数の品詞がある場合は、文章で使用されている品詞を基に分類を行った。計算方法は、 $(GL \text{ (General Links: 本文に3回以上登場する内容語)} / \text{選択肢の総語数}) + (UL \text{ (Unique Links: 本文に1, 2回登場する内容語)} / \text{選択肢の総語数}) = \text{パラフレージング・レベル}$ とした(政所, 2018を参照)。合計が50%以上の場合はNear Copy (NC), 20～49%の場合はMinimal Revision (MIR), 1～19%の場合はModerate Revision (MOR), 0%の場合はSubstantial Revision (SR)として分類を行った(表2参照)。

■表2: パラフレージング・レベルの分類

| パラフレージング・レベル              | UL, GLの語数の合計が総語数に占める割合 |
|---------------------------|------------------------|
| Near Copy (NC)            | 50%以上                  |
| Minimal Revision (MIR)    | 20～49%                 |
| Moderate Revision (MOR)   | 1～19%                  |
| Substantial Revision (SR) | 0%                     |

(注) 政所(2018)p. 89表2を参照

## 2.4 生成AIでの選択肢作成

ChatGPT-4oを使用し、正答選択肢のパラフレーズを行った。使用したプロンプトは「(正答選択肢を入れる) Create three paraphrased sentences of the above sentences using CEFR B1 level vocabulary.」である。本文を入力して正答選択肢が本文の語を含む割合を調整するプロンプトも検討したが、当初は生成AIは数字を正確に扱うことができないと言われていたため、このようなプロンプトを使用した。また、生成AIに馴染みのない教員も手軽に使うことができるよう、プロンプトの指定は簡素にしている。なお、事前に一つの文章(3項目)の正答選択肢について、前述のプロンプトで3つのパラフレーズ案を作成し、本文の内容語を含む割合を比較・検討し、十分にパラフレーズされることが確認できたため、プロンプトでは「3つの選択肢」としている。このプロンプトでは、主語や他の錯乱肢にはない部分を自動的に補ってしまうという課題があったが、それらを削除してそのまま使用できる場合や、文頭にToを補えば使用できる場合はそれらの選択肢を採用した(例:It was noticed that bonobos…とパラフレーズが出力された場合、錯乱肢がbonobosから始まっている場合は、It was noticed that の部分を削除したbonobosから始まる文章の選択肢として採用する)。しかし、文法的に修正が必要な場合、語彙を変えなければならない場合は、その選択肢は採用しなかった(1:主語が三単現であり、助動詞(will)から始まる選択肢をパラフレーズした際に、主語が補われてI am going to…になっている場合は… is going to…のように語彙を変えなければならないもの(項目1)。2:主語が複数形の場合にwas not enough…となっており、were not enough…のように語彙を変えなければならないもの(項目18))。

これらの採用しなかった2つの選択肢以外の全ての選択肢について、パラフレーズ・レベルと本文の内容語を使用している割合(%)を計算し、本文の内容語を使用している割合(%)が最も低いものを採用した。選択肢の総語数と内容語の種類(people, English等の名詞)によっては、パラフレーズする語彙には限界があること、また、短い選択肢は本文の内容語を使用した割合が高くなりやすいため、本文の内容語を使用している割合がオリジナルの正答選択肢よりも高い場合があった。その場合には、もう一度出力させると作成者側の手間が増えてしまい、研究目的の一つである「選択肢を効率的に作成すること」を達成できなくなってしまうことから、パラフレーズされた選択肢のうち最も本文の内容語が含まれている割合が低いものを使用した。

■表3: 生成AIを使用して言い換えをした選択肢の一例

|         | GL | UL | 総語数 | パラフレーズ・レベル |
|---------|----|----|-----|------------|
| 出力1     | 3  | 0  | 9   | 33.3%      |
| 出力2     | 4  | 2  | 10  | 60.0%      |
| 出力3(採用) | 3  | 0  | 11  | 27.3%      |

オリジナルの正答選択肢:Keep an eye on how words are used in the English language.

- 出力1 :Pay attention to how words are used in English.
- 出力2 :Watch carefully how people use words in the English language.
- 出力3(採用) :Notice how words are used when speaking or writing in English.

参考までに、オリジナルテストの選択肢のパラフレーズの度合いを調べたところ、表4, 5の通りになった。合計で28項目使用し、選択肢は各問4つのため、選択肢の合計は112だった。政所(2018)では、選択肢の内容語のうち、本文で使用されている割合を算出していたため、その方法を採用した(2.3 参照)。

■表4: オリジナルテストの全ての選択肢におけるパラフレージング・レベル

|     | 2023年度第3回 | 2024年度第1回 | 2024年度第2回 | 合計  | 割合    |
|-----|-----------|-----------|-----------|-----|-------|
| NC  | 10        | 6         | 4         | 20  | 17.9% |
| MIR | 29        | 19        | 23        | 71  | 63.4% |
| MOR | 7         | 4         | 4         | 15  | 13.4% |
| SR  | 2         | 3         | 1         | 6   | 5.4%  |
| 合計  | 48        | 32        | 32        | 112 | 100%  |

■表5: オリジナルテストの正答選択肢におけるパラフレージング・レベル

|     | 2023年度第3回 | 2024年度第1回 | 2024年度第2回 | 合計 | 割合    |
|-----|-----------|-----------|-----------|----|-------|
| NC  | 3         | 0         | 2         | 5  | 17.9% |
| MIR | 8         | 7         | 5         | 20 | 71.4% |
| MOR | 1         | 0         | 1         | 2  | 7.1%  |
| SR  | 0         | 1         | 0         | 1  | 3.6%  |
| 合計  | 12        | 8         | 8         | 28 | 100%  |

オリジナルテストの全ての選択肢のパラフレージングレベル(表4参照)と正答選択肢のパラフレージング・レベル(表5参照)は、それぞれMIRが最も多く、次にNCが多いという政所(2018)の結果と一致した。政所(2018)では、清水(2005)で分類されている質問の一つであるパラフレーズ質問のみに焦点を当てている。本研究では質問の種類での分類は行っていないが、全体的に同様のパラフレージング・レベルの傾向があることが確認できた。

また、事前に英語教育を専攻している大学院生2名に正答選択肢をパラフレーズしたテスト問題を解いてもらい、選択肢に不自然な点がないかの点検と解答時間の記録を依頼した。選択肢では、メールを読み取る形式の問題で、To give her your best wishes for…と出力されていた選択肢において、メールの相手ではなく読み手を指すyourが補われていた部分について指摘を受けたため、yourを削除して使用した。解答時間については、依頼した2名は全体を通して30分から40分だったが、英語を専門としていない人も調査の対象とするため余裕を持って1時間で設定した。

#### 2.4.1 調査協力者

協力者は国立大学・大学院に所属する学生65名であり、所属学部は教育学部・工学部・理学部・国際教養学部・園芸学部・法政経学部等、多岐にわたる。

#### 2.4.2 実施方法

2025年1月中旬から2月上旬にかけてテストを実施した。大学関係者のみにアクセスを制限した上でGoogle Formで参加者を募り、協力者の希望時間に合わせて16グループに分けて実施した。テスト実施前に、研究参加は任意であり途中退室も可能であることや、データは統計的に処理し個人は特定できない旨を記した書面を配布した。その書面を実施者が読み上げて説明をし、参加者の協力への同意を署名を以って確認した。今回は研究テーマや研究目的を事前に伝えると結果に影響を及ぼす可能性があるため、事前の同意書の段階では研究テーマや研究目的を伝えていない。また、テストには匿名での参加が可能である

点を伝えたところ、全員が匿名での参加を選んだ。

制限時間は1時間とし、オリジナル(選択肢そのままのもの)、パラフレーズ(正答の選択肢のみ生成AIでパラフレーズしたもの)の2種類を使用し、7つの文章題でカウンターバランスをとった(表6参照)。テストの表紙には、全ての問題を解き切ること、解答を書いたら1項目ずつ自信度に丸をつけること、提示された順番通りに問題に解答すること、の3点について守るように依頼する文章を載せており、テスト前にそれらの説明が書かれた表紙を見せながら実施者が読み上げて説明を行った。英文のジャンルが2種類あり、メールの文章が3つと説明文が4つであること、文章題一つにつき7分くらいの目安で行えば1時間以内で終了することを事前に説明した。事前に問題の全ての分量や内容を確認することは可能であるとし、解く際には前から順番に解くことを徹底してもらうよう伝えた。同時に1名から8名でテストを行った。同時に参加した協力者が1〜2名であり、かつ全員が1時間以内で解き終えた場合は制限時間の1時間を待たずに終了した。

■表6: 問題タイプ別受験者数

| Type | 問題順番          | オリジナル   | パラフレーズ |
|------|---------------|---------|--------|
| 1    | A→B→C→D→E→F→G | 4       | 5      |
| 2    | B→C→D→E→F→G→A | 5       | 5      |
| 3    | C→D→E→F→G→A→B | 5       | 5      |
| 4    | D→E→F→G→A→B→C | 5       | 5      |
| 5    | E→F→G→A→B→C→D | 5       | 4      |
| 6    | F→G→A→B→C→D→E | 4       | 4      |
| 7    | G→A→B→C→D→E→F | 4       | 4      |
|      |               | 32名(注1) | 32名    |

(注1) 時間内に解答ができなかった1名を除いた数

2.4.3 分析方法

オリジナル(選択肢そのままのもの)を解いた人数は33名、パラフレーズ(正答の選択肢のみ生成AIでパラフレーズしたもの)を解いた人数は32名だった。このうち、オリジナルの問題を解いた1名は制限時間内での解答ができなかったため、分析から除外した。また、パラフレーズを解いた人のうち、1名が1項目について解答なし、1名が1項目について解答はしているものの自信度の解答がなかった。前者は問題用紙に下線が引いてあり、読んだ形跡があった。この場合、実際のテストでもわからない問題はとばすという方略があるため、質問を誤って飛ばしたのではなく意図的に飛ばしたとみなし誤りとして扱った。分析は、正答率と自信度のそれぞれにおいて対応のないt検定を行った。なお、今回は正答率と自信度の関係については分析対象としていない。

3 結果と考察

R(R Core Team, 2025)を使用して信頼性係数を出したところ、オリジナルは  $\alpha = .85$ 、パラフレーズは  $\alpha = .87$  であり、どちらも高い信頼性が確認できた。

3.1 パラフレーズした選択肢

オリジナルとパラフレーズのそれぞれの正答選択肢における本文の語を含む割合を一覧にして以下に示す。

■表7: 2023年度第3回実施分

| 項目番号   | A     |       |       | B     |       |       |       | C     |       |       |       |       |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|        | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | 11    | 12    |
| オリジナル  | 36.4% | 10.0% | 50.0% | 33.3% | 23.1% | 22.2% | 30.8% | 20.0% | 55.6% | 23.1% | 54.5% | 42.9% |
| パラフレーズ | 22.2% | 7.1%  | 33.3% | 27.3% | 25.0% | 0%    | 20.0% | 12.5% | 30.0% | 18.2% | 41.7% | 35.7% |

■表8: 2024年度第1回実施分

| 項目番号   | D     |    |       | E     |       |       |       |       |
|--------|-------|----|-------|-------|-------|-------|-------|-------|
|        | 13    | 14 | 15    | 16    | 17    | 18    | 19    | 20    |
| オリジナル  | 37.5% | 0% | 22.2% | 30.8% | 28.6% | 33.3% | 37.5% | 38.5% |
| パラフレーズ | 14.3% | 0% | 16.7% | 18.2% | 18.2% | 25.0% | 0%    | 28.6% |

■表9: 2024年度第2回実施分

| 項目番号   | F     |       |       | G     |       |       |       |       |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|
|        | 21    | 22    | 23    | 24    | 25    | 26    | 27    | 28    |
| オリジナル  | 35.7% | 30.0% | 14.3% | 30.8% | 40.0% | 66.7% | 30.8% | 60.0% |
| パラフレーズ | 30.8% | 18.2% | 12.5% | 18.2% | 33.3% | 27.3% | 8.3%  | 33.3% |

■表10: 本文の語を使用した割合ごとの選択肢分布

| 割合      | オリジナル                            |     | パラフレーズ                    |     |
|---------|----------------------------------|-----|---------------------------|-----|
|         | 項目番号                             | 項目数 | 項目番号                      | 項目数 |
| 100%    |                                  | 0   |                           | 0   |
| 99%~90% |                                  | 0   |                           | 0   |
| 89%~80% |                                  | 0   |                           | 0   |
| 79%~70% |                                  | 0   |                           | 0   |
| 69%~60% | 26,28                            | 2   |                           | 0   |
| 59%~50% | 3,9,11                           | 3   |                           | 0   |
| 49%~40% | 12,25                            | 2   | 11                        | 1   |
| 39%~30% | 1,4,7,13,16,18,19,20,21,22,24,27 | 11  | 3,9,12,21,25,28           | 7   |
| 29%~20% | 5,6,8,10,15,17                   | 7   | 1,4,5,7,18,20,26          | 6   |
| 19%~10% | 2,23                             | 2   | 8,10,13,15,16,17,22,23,24 | 9   |
| 9~1%    |                                  | 0   | 2,27                      | 2   |
| 0%      | 14                               | 1   | 6,14,19                   | 3   |
|         |                                  | 28  |                           | 28  |

表7～表10のデータから、項目5、項目14の合計2つを除いてパラフレーズの方が正答選択肢における本文の語を含む割合が低くなっていることが明らかになった。項目14についてはオリジナルの正答選択肢が本文の語を含む割合が0%だったため、それ以上に本文の語を含む割合を下げることはできなかった。項目5についてはオリジナルとパラフレーズのそれぞれの正答選択肢に対して、本文の語はそれぞれ3語ずつ含まれており、オリジナルの選択肢の総語数が13語に対し、パラフレーズの総語数が12語と少なかったことが、パラフレーズの正答選択肢における本文の語を含む割合が高くなった原因である。項目5の正

答選択肢を以下に示す。

オリジナル : a large number of people are using the word in the same way. (13語)

パラフレーズ: a lot of people are using the word in a similar way. (12語)

オリジナルでは people, word, same の3語が本文の語を使用しており, パラフレーズでは lot, people, word の3語が本文の語を使用している。

### 3.2 テスト受験協力者の外部英語試験のスコア

2種類のテストは正答選択肢のみ異なるが問題文が同じであり, 同じ人が2種類のテストを解くことはできないため, オリジナル, パラフレーズを解いたそれぞれの集団が同質となるように最新の英語の資格試験のスコアを事前に聞いた上で割り振った(表11参照)。

英検以外の外部試験の場合, 4技能を測っているもの (IELTS, TOEFL 等) であれば CEFR 対応表を参照して対応する英検の級とみなした。また, 例外として不合格のスコアが書かれていた場合, その記載されたスコアに相当する級として扱った(例: 英検準1級不合格スコア2100点は英検2級相当として扱う)。

TOEIC はリーディング, リスニングの2技能のみのスコアだったため, 英検の取得級やその他の4技能を測るテストのスコアとは別に, 2つのグループで均等になるように分けた(表12参照)。

■表11: 外部英語試験のスコア別の人数

|        | 英検1級相当 | 英検準1級相当 | 英検2級相当 | 英検準2級相当 | TOEIC | 取得歴なし | 計  |
|--------|--------|---------|--------|---------|-------|-------|----|
| オリジナル  | 0      | 13      | 7      | 0       | 6     | 6     | 32 |
| パラフレーズ | 1      | 12      | 8      | 1       | 5     | 5     | 32 |

■表12: TOEICスコアの得点分布(注:外部試験の受験経験がTOEICのみの協力者)

| 点数     | 500～ | 600～ | 700～ | 800～ | 合計 |
|--------|------|------|------|------|----|
| オリジナル  | 2    | 1    | 2    | 1    | 6名 |
| パラフレーズ | 2    | 2    | 0    | 1    | 5名 |

### 3.3 得点

問題は合計28項目使用し, 正答の選択肢を選ぶことができていた場合1点, 誤りの選択肢を選んだ場合は0点とし, 全部で28点満点として処理した。表13のテストの得点の記述統計を参照すると, どちらも天井効果が見られた ( $M+SD$ : オリジナル28.82, パラフレーズ28.20)。テスト受験協力者が最近取得した英語資格のアンケートからも, 全体的に英語への関心や意識が高いことが影響していると考えられる。

■表13: テストの得点

|        | <i>N</i> | <i>M</i> | <i>SD</i> | <i>Max</i> | <i>Min</i> |
|--------|----------|----------|-----------|------------|------------|
| オリジナル  | 32       | 25.56    | 3.25      | 28         | 16         |
| パラフレーズ | 32       | 23.78    | 4.41      | 28         | 11         |

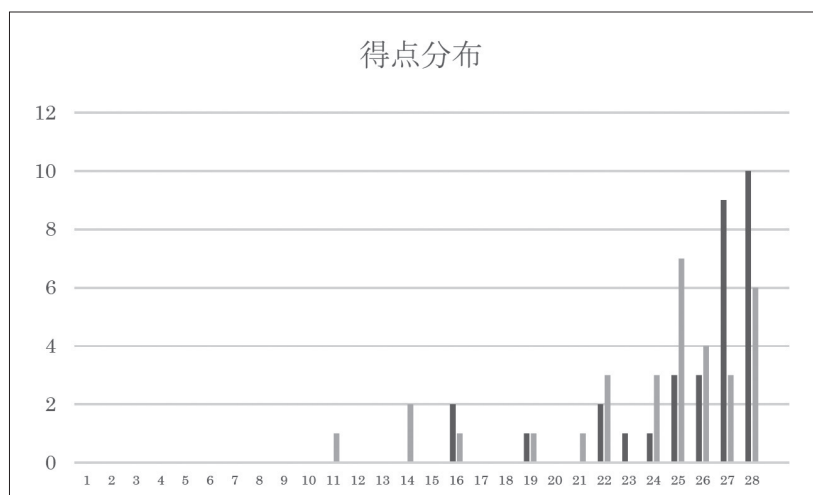
(注) 満点は28点

オリジナルとパラフレーズの受験者の得点において対応のない  $t$  検定を行ったところ, 有意差がなかった ( $t = 1.81, df = 62, p = .076$ )。一方, 項目正答率についても同様に対応のない  $t$  検定を行ったところ, 有意差があった ( $t = 2.56, df = 54, p = .013$ )。表24に記載した項目正答率の平均では, オリジナルでは91.3%

に対して、パラフレーズでは84.9%であり、全体的にパラフレーズの項目の難易度が高いことが示された。

■表14: 得点分布

|        | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 合計 |
|--------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| オリジナル  | 0  | 0  | 0  | 0  | 0  | 2  | 0  | 0  | 1  | 0  | 0  | 2  | 1  | 1  | 3  | 3  | 9  | 10 | 32 |
| パラフレーズ | 1  | 0  | 0  | 2  | 0  | 1  | 0  | 0  | 1  | 0  | 1  | 3  | 0  | 3  | 7  | 4  | 3  | 6  | 32 |



■図1: 得点分布

(注)濃:オリジナル, 薄:パラフレーズ

全28項目のうち、オリジナルを解いた受験者全員が正解だった項目が、項目6、項目9、項目10、項目14、項目24、項目26、項目28の合計7項目だった。パラフレーズを解いた受験者全員が正解だった項目は、項目9、項目24、項目25の合計3項目だった。オリジナルとパラフレーズの両方で全員が正解した項目は項目9、項目24であり、項目25のみがパラフレーズのみで受験者全員が正解だった。項目25はオリジナルのテストでも誤りを選んだ人は一人のみだったため、項目25においてはオリジナルとパラフレーズの両者の項目において、大きな差はないとみなすことができる。

オリジナルとパラフレーズの項目正答率について対応のない $t$ 検定をしたところ、 $p$ 値が.05未満の項目は合計2つだった(表15参照)。その2つにおいて、オリジナルの方がパラフレーズよりも項目正答率が有意に高い結果がでた。全体としてパラフレーズの方がオリジナルよりもやや難しい傾向があると言えることから、正答選択肢に含まれる本文の語の割合がより低い場合については難易度がやや上がることがわかる。

■表15: 項目正答率に有意差がある項目

| 項目番号 | 項目正答率 |        | 2種類のテストの正答率の差(%)<br>(オリジナルーパラフレーズ) | $p$ 値 |
|------|-------|--------|------------------------------------|-------|
|      | オリジナル | パラフレーズ |                                    |       |
| 1    | 93.8  | 75.0   | 18.8                               | .039  |
| 22   | 84.4  | 62.3   | 22.1                               | .049  |

全体的に受験者の正答率が高く天井効果があり、オリジナルとパラフレーズで正答選択肢の選択率にも大きな差はなかったため、今回は各選択肢の選択率については分析対象外としている。オリジナルとパラフレーズで、異なる錯乱肢を選んだ人数が5人以上異なる項目を探したところ、項目12が該当したため、項目12における正答選択率と錯乱肢との関係については表16に示す。

■表16: 項目12 選択肢別の解答詳細

| 選択肢    | オリジナル     |          |      | パラフレーズ    |          |      |
|--------|-----------|----------|------|-----------|----------|------|
|        | 本文の語を含む割合 | 選択した受験者数 | 割合   | 本文の語を含む割合 | 選択した受験者数 | 割合   |
| 1      | 42.9      | 3        | 9.4  | -         | 1        | 3.1  |
| 2 (正答) | 42.9      | 25       | 78.1 | 35.7      | 23       | 71.9 |
| 3      | 35.7      | 4        | 12.5 | -         | 2        | 6.3  |
| 4      | 46.2      | 0        | 0    | -         | 6        | 18.8 |

(注)-はオリジナルと同じ値を表す

#### 項目12

問題文: Which of the following statements is true?

錯乱肢: 1. Marie Curie had a lot of difficulty remembering facts when she was a child.

2. (正答選択肢)

3. Marie Curie was taught science by a female professor at a university in Paris.

4. Marie Curie wanted to study science after finding out how X-ray machines worked.

オリジナルの正答選択肢

3. Marie Curie and her sister agreed to help each other to study at university.

パラフレーズの正答選択肢

3. Marie Curie and her sister promised to assist each other in their university education.

項目12の正答選択肢は2であり、オリジナルで正答選択肢を選んだ割合は78.1%、パラフレーズで正答選択肢を選んだ割合は71.9%だった。錯乱肢4は、オリジナルでは選択した人が0名だったのに対しパラフレーズでは6名いた。オリジナルの正答選択肢の本文の語を含む割合が42.9%であり、パラフレーズの正答選択肢の本文の語を含む割合は35.7%だったのに対して、錯乱肢4の本文の語を含む割合は46.2%だった。このことから、オリジナルの正答選択肢と錯乱肢4の本文の語を含む割合には大きな差はなかったが、パラフレーズの正答選択肢よりも錯乱肢4の本文の語を含む割合には10.5パーセンテージポイントの差があり、本文の語を含む割合が高い錯乱肢4の方を選んでしまった可能性がある。

■表17: 知識及び技能を測る項目と思考力・判断力・表現力等を測る項目の平均正答率

|         | オリジナル            |                        | パラフレーズ           |                        |
|---------|------------------|------------------------|------------------|------------------------|
|         | 知識及び技能<br>(16項目) | 思考力・判断力・表現力等<br>(12項目) | 知識及び技能<br>(16項目) | 思考力・判断力・表現力等<br>(12項目) |
| 平均項目正答率 | 93.4             | 88.5                   | 84.8             | 85.2                   |

知識及び技能を測る項目と思考力・判断力・表現力等を測る項目について、対応のない $t$ 検定を行った結果、オリジナルとパラフレーズともに有意差はなかった(オリジナル: $t = 1.5587$ ,  $df = 26$ ,  $p = .131$ , パラフレーズ: $t = -0.098693$ ,  $df = 26$ ,  $p = 0.922$ )。項目正答率においては両者に違いは確認できなかった。

### 3.4 自信度

自信度は項目ごとに低中高の3段階で丸をつける方法を採用した。協力者には3段階の自信度の評価を28項目全て解答してもらい、高を3、中を2、低を1として計算を行ったため、一人当たりの最大値は84になる。表18には、オリジナルとパラフレーズのそれぞれの自信度における記述統計を示す。オリジナルとパラフレーズにおける全ての項目の自信度を対応のない $t$ 検定をしたところ、有意差がなかった( $t = .81$ ,  $df = 60$ ,  $p = .42$ )。問題ごとに有意差を確認したところ、表19の通りになった。

■表18: 2種のテストの自信度

|        | <i>N</i> | <i>M</i> | <i>SD</i> | <i>Max</i> | <i>Min</i> |
|--------|----------|----------|-----------|------------|------------|
| オリジナル  | 32       | 69.13    | 11.65     | 83         | 37         |
| パラフレーズ | 30       | 66.53    | 13.24     | 83         | 32         |

(注) 全28問を自信度3段階で解答し、一人当たりの最大値は84

■表19: 自信度に有意差がある項目

| 項目番号 | 自信度 平均 |        | 2種類のテストの差<br>(オリジナルーパラフレーズ) | <i>p</i> 値 |
|------|--------|--------|-----------------------------|------------|
|      | オリジナル  | パラフレーズ |                             |            |
| 1    | 2.0    | 2.4    | -0.4                        | .035       |
| 2    | 2.3    | 1.8    | 0.5                         | .013       |
| 6    | 2.8    | 2.3    | 0.5                         | .0038      |
| 13   | 2.6    | 2.2    | 0.4                         | .045       |
| 17   | 2.1    | 2.5    | -0.4                        | .023       |
| 26   | 2.7    | 2.3    | 0.4                         | .025       |

(注) 自信度の平均は高=3、中=2、低=1として算出

テスト全体として自信度に有意な差は確認できなかったが、項目ごとの自信度について $t$ 検定を行った結果、有意差の確認ができた項目は合計6つだった。項目17は、オリジナルとパラフレーズの正答率が同じであり87.5%だったものの(項目正答率については表24参照)、自信度の平均はパラフレーズの方が高かった。項目1と項目17を除き、本文の語を含む割合が低い正答選択肢を使用したパラフレーズの問題の方が自信度が低く、その結果に有意差が確認できた。項目1と項目17の具体的な項目と選択肢を以下に示す。自信度について有意差がある項目6項目について、項目1,13の2つは「知識及び技能」を測る項目であり、項目2,6,17,26の4つは「思考力、判断力、表現力等」を測る問題だった。全体の項目の傾向として「知識及び技能」を測る項目が16項目に対して「思考力、判断力、表現力等」を測る項目が12項目であり、「思考力、判断力、表現力等」を測る項目の方が全体に占める数が少ない。しかし、自信度について有意差のある項目6つのうち「思考力、判断力、表現力等」を測る項目は4つあり、「知識及び技能」を測る項目よりも、6つのうちに占める割合が多い。このことから、「思考力、判断力、表現力等」を測る項目に対する自信度は相対的に下がる可能性があると言える。

項目1 知識及び技能を測る問題

問題文: Some employees from the IT department

- 錯乱肢: 1. have been selected to work at several job fairs at different colleges.  
2. were asked to give a guest lecture to computer engineering students.  
3. (正答選択肢)  
4. are being sent to another department to help train new workers.

オリジナルの正答選択肢

3. will take part in an event at a university in Petersburg.

パラフレーズの正答選択肢

3. will join an event at a university in Petersburg.

項目1の正答選択肢は、動詞のみ異なる(オリジナル:take part in, パラフレーズ:join)。オリジナルの方が自信度が有意に低い点については、オリジナルを解いた人のうち誤りの選択肢である2を選んだ2名が自信度を低で選択している。また、正答選択肢を選んだ30名中10名が自信度を低で選択し、加えて9名が自信度を中で選択している(表20, 21を参照)。オリジナルの方が自信度が低い理由として、オリジナルで使用されていたtake part inはコロケーションであったことが挙げられる。加えて、パラフレーズで使用されていたjoinの方がオリジナルで使用されていたtake part inよりも、受験者にとって馴染みのある語彙であったことが影響していると考えられる。

■表20: 項目1の選択肢を選んだ人数

| 選択肢   | 選択肢を選んだ人数 |        |
|-------|-----------|--------|
|       | オリジナル     | パラフレーズ |
| 1     | 0         | 3      |
| 2     | 2         | 4      |
| 3(正答) | 30        | 24     |
| 4     | 0         | 1      |
| 合計    | 32        | 32     |

■表21: 項目1の自信度の人数分布

| 自信度 | 選んだ人数 |        |
|-----|-------|--------|
|     | オリジナル | パラフレーズ |
| 低   | 12    | 5      |
| 中   | 9     | 9      |
| 高   | 11    | 18     |
| 合計  | 32    | 32     |

## 項目17 思考力, 判断力, 表現力等を測る問題

問題文: What is one thing the members of the Arts and Crafts movement believed?

錯乱肢: 1. Machines would never be able to make objects as quickly as artists could.

2. Paintings and sculptures were not worth making because they were not useful.

3. (正答選択肢)

4. Designers should specialize in producing just one or two types of items.

オリジナルの正答選択肢

3. The lives of people would improve if they returned to making things by hand.

パラフレーズの正答選択肢

3. If people returned to handmade production, their lives would become better.

項目17は、本文を2文以上参照した上で、推論する必要である項目であるため、「思考力, 判断力, 表現力等」を測る項目に分類した(分類方法は2.2を参照)。総語数は、オリジナルが14語でありパラフレーズが11語だった。オリジナルは結論を先に示したのちに条件(if)を示しているのに対し、パラフレーズは条件(if)を示したのちに結論を示しており、パラフレーズの方が日本語の文章の順番に近い形であることが、自信度が高くなった要因として予想できる。Oxford Advanced Learner's Dictionary 第10版によると、オリジナルの正答選択肢に含まれているby handには、(1)機械ではなく人の手によるもの、(2)手書きや手渡しであること、の2つの意味が記載されている。このように、by handには複数の意味があることがオリジナルの正答選択肢への自信度を下げている要因であると考えられる。また、by handをWeb Para Newsというコーパス検索サイトで調べると、byのヒットは500件(以上)、handのヒットは178件ある中で、by handのコロケーションのヒットは1件のみだった。別のコーパスサイトとして、小学校・中学校・高等学校の教科書における語彙を検索することができるTeachLex(Sato, 2025)を使用し、handとby両方の登場頻度とコロケーションを調べた。handの登場頻度は、全ての教科書を合計して小学校で39回、中学校で27回であり、byの登場頻度は、小学校で107回、中学校で178回だった。語彙同士の結びつきを表しているワードクラウドでも、byとhandの強い結びつきは小学校・中学校・高等学校のいずれの学校種においても確認することができなかった。以上のことから、by handのコロケーションの使用頻度が低く、自信度を下げている要因として考えられる。一方で、パラフレーズの正答選択肢に使用されているhandmadeという語は日本語でもよく使用されていることから、パラフレーズの方が本文の語を含む割合が低いにもかかわらず、自信度が高くなったと考えられる。

■表22: 項目17の選択肢を選んだ人数

| 選択肢   | 選択肢を選んだ人数 |        |
|-------|-----------|--------|
|       | オリジナル     | パラフレーズ |
| 1     | 2         | 2      |
| 2     | 1         | 0      |
| 3(正答) | 28        | 28     |
| 4     | 1         | 2      |
| 合計    | 32        | 32     |

■表23: 項目17の自信度の人数分布

| 自信度 | 選んだ人数 |        |
|-----|-------|--------|
|     | オリジナル | パラフレーズ |
| 低   | 9     | 3      |
| 中   | 11    | 9      |
| 高   | 12    | 20     |
| 合計  | 32    | 32     |

### 3.5 弁別力

テスト全体の得点と、その項目の正答率の相関を表す点双列相関係数を求めた(表24を参照)。竹内・水本(2023)を参照し、項目一全体得点相関係数である点双列相関係数は0.30以上であれば能力の弁別ができていて項目として結果を解釈した。

オリジナルは7項目について全員が正解、パラフレーズは3項目について全員正解の項目があったため、それらは分析の過程で除外されているが、それ以外の項目の弁別力は表24の通りである。項目1, 項目11, 項目13, 項目17, 項目18, 項目20, 項目23の計7項目については、オリジナルよりもパラフレーズの方が.10以上高い値が出ており、項目2, 項目4, 項目16, 項目22, 項目27の計5項目については、パラフレーズよりもオリジナルの方が.10以上高い値が出ている。どちらのテストも概ね弁別力は高く、弁別力が.30未満の項目は、オリジナルでは項目11, 項目13, 項目15, 項目18の計4項目、パラフレーズでは項目4, 項目6, 項目8, 項目16の計4項目だった。また、オリジナルの各項目における弁別力の平均値は、.50に対して、パラフレーズの各項目における弁別力の平均値は.49であり、差はほとんどなかった。この結果から、オリジナルとパラフレーズで、弁別力に大きな差は見られず、弁別力はどちらも同等であることが明らかになった。

■表24: 項目別正答率・弁別力・自信度

| 項目 | 項目別正答率 |        | 弁別力   |        | 自信度   |        |
|----|--------|--------|-------|--------|-------|--------|
|    | オリジナル  | パラフレーズ | オリジナル | パラフレーズ | オリジナル | パラフレーズ |
| 1  | 93.8   | 75.0   | .52   | .67    | 2.0   | 2.4    |
| 2  | 71.9   | 65.6   | .51   | .41    | 2.3   | 1.8    |
| 3  | 90.6   | 84.4   | .55   | .54    | 2.5   | 2.3    |
| 4  | 90.6   | 87.5   | .85   | .25    | 2.6   | 2.2    |
| 5  | 96.9   | 87.5   | .53   | .62    | 2.7   | 2.7    |
| 6  | 100.0  | 87.5   | -     | .26    | 2.8   | 2.3    |
| 7  | 81.3   | 87.5   | .65   | .56    | 2.4   | 2.5    |
| 8  | 87.5   | 93.8   | .67   | .075   | 2.8   | 2.7    |
| 9  | 100.0  | 100.0  | -     | -      | 2.9   | 2.8    |
| 10 | 100.0  | 90.6   | -     | .37    | 2.7   | 2.6    |
| 11 | 96.9   | 90.6   | .20   | .62    | 2.7   | 2.5    |
| 12 | 78.1   | 71.9   | .53   | .55    | 2.1   | 1.9    |
| 13 | 93.8   | 78.1   | .24   | .54    | 2.6   | 2.2    |
| 14 | 100.0  | 93.8   | -     | .57    | 2.4   | 2.6    |
| 15 | 96.9   | 96.9   | .031  | .40    | 2.6   | 2.5    |
| 16 | 87.5   | 90.6   | .47   | .28    | 2.5   | 2.6    |
| 17 | 87.5   | 87.5   | .53   | .69    | 2.1   | 2.5    |
| 18 | 87.5   | 71.9   | .27   | .40    | 2.3   | 1.9    |
| 19 | 84.4   | 78.1   | .55   | .44    | 2.2   | 2.0    |
| 20 | 90.6   | 81.3   | .42   | .77    | 2.4   | 2.4    |
| 21 | 93.8   | 78.1   | .48   | .47    | 2.4   | 2.1    |
| 22 | 84.4   | 62.5   | .87   | .52    | 2.0   | 2.3    |
| 23 | 71.9   | 75.0   | .36   | .56    | 2.5   | 2.5    |
| 24 | 100.0  | 100.0  | -     | -      | 2.9   | 2.9    |
| 25 | 96.9   | 100.0  | .53   | -      | 2.4   | 2.4    |
| 26 | 100.0  | 90.6   | -     | .52    | 2.7   | 2.2    |
| 27 | 93.8   | 93.8   | .76   | .66    | 2.4   | 2.4    |
| 28 | 100.0  | 84.4   | -     | .56    | 2.4   | 2.2    |
| 平均 | 91.3   | 84.9   | .50   | .49    | 2.5   | 2.4    |
| SD | 8.16   | 9.99   | .20   | .16    | 0.26  | 0.27   |

(注) 弁別力における - は受験者全員が正解だったため、除外された問題

## 4 結論

本研究は、正答選択肢における本文の内容語を言い換えた割合を調整した場合、項目正答率・テスト受験者の正答選択肢に対する自信度・弁別力のそれぞれがどのように変化するかを検証することを目的として行った。また、教育現場での読解問題の作成において、客観的な採点が可能であり、実用性の面で採点が容易であるという特徴を持つ多肢選択式テストに焦点を当て、生成 AI を活用した選択肢の作成方法を検討した。ChatGPT-4o を使用して本文の語を含む割合を調整した正答選択肢を作成し、2種類のテストを合計64名の大学生・大学院生を対象に実施した。2種類のテストはともに信頼性は高く、本文の内容語を含む割合を低くした正答選択肢を使用しても、オリジナルの項目と同等の高い信頼性が得られることが明らかになった。本文の語を使用する割合がより低い正答選択肢を含む項目と、本文の語を使用する割合がより高い正答選択肢を含む項目とを比較した結果について、先に設定した3つの研究課題の結論を以下に示す。

研究課題1の「項目正答率がどの程度変化するか明らかにすること」について、本文の語を使用する割合を低くした正答選択肢を使用した場合、項目正答率が有意に低くなり、全体的に難易度がやや高くなるという結果が明らかになった。このことから、項目正答率は下がるという仮説は支持された。政所(2018)では、特に低位群では、正答選択肢が本文の語をあまり含まない場合、より本文の語を含む錯乱肢を選ぶ傾向があることが明らかにされている。本研究では天井効果が見られたため能力別の比較はできなかったものの、本文の語を使用する割合が選択肢の選ばれやすさに影響を及ぼす可能性があると言える。

研究課題2の「受験者の項目に対する自信度がどの程度変化するか明らかにすること」について、自信度は下がるという仮説は部分的に支持された。本文の語を含む割合が低い正答選択肢を使用したパラフレーズの方が自信度は低く、有意差がある項目は28項目中6項目だった。自信度についても、本文の語を含む割合が低い場合には、表面的に難しく見える可能性があり、自信度が下がるという仮説は部分的に支持できると言える。

研究課題3の「受験者の能力を識別する弁別力はどの程度変化するか明らかにすること」について、弁別力においてはほぼ同等であることが示唆された。そのため、弁別力は上がるという仮説は支持されなかった。

本研究では天井効果が見られたが、オリジナルテストの選択肢は信頼性・項目弁別力ともに十分であり、正答選択肢をパラフレーズしてもオリジナルテストと同等の信頼性や弁別力が保たれることが示唆された。生成 AI を使用してパラフレーズをしても語彙の難易度が著しく上がることもなく、錯乱肢と比較しても違和感のない選択肢を作成することができたと言える。一方で、本研究の結果からは、傾向を確認するための量的なデータを十分に提示することができず、選ばれやすさの要素が問題によって異なる場合があることも示された。そのため、どのような問題で本文の語を含む割合が選ばれやすさの要素になるかについて、さらに研究を行う必要がある。テスト作成で生成 AI を用いて言い換えをする作成方法については、テストに天井効果が見られただけでなく、本研究の結果のみでは一般化するために十分な量的なデータを提示することはできなかったことから、一部支持できると言える。生成 AI を積極的に使用することで、実用性の観点から作成時間も短縮することができる。教員による最終チェックは必要だが、業務の効率化につなげることができるだろう。

最後に、本研究の限界点として、天井効果が見られた点と、語彙レベルの調整の2点が挙げられる。使用した2種類のテストでは受験者の得点が高く、どちらも天井効果が見られたため能力別の傾向を明らかにすることができなかった。語彙レベルについては、英検2級の問題では言い換えができる語彙の量が十分だったため、生成 AI を使用して言い換えた正答選択肢の語彙が著しく難しくなることはなかった。しかし、言い換えをする際に使用できる語彙の量が少ない初級レベルの学習者向けのテストでは、本研究とは異なる結果が出ることが予想できる。今後は生成 AI を用いて作成したテストの妥当性と信頼性について、

読解の本文や錯乱肢の言い換え、異なる語彙レベルでの言い換えをしたテストを用いて検証したい。

## 謝辞

本研究の実施の機会を与えてくださった公益財団法人 日本英語検定協会の皆さま、選考委員の先生方に厚く御礼申し上げます。特に、本研究を担当してくださいました齊田智里先生には貴重なご助言を賜りましたことを深く感謝申し上げます。また、千葉大学の星野由子先生には、本研究に関して計画・実施・報告書の作成に至るまで親身にご指導いただきましたこと、心から感謝申し上げます。最後に、本調査を実施するにあたり、問題の分類やテストの事前チェックにご協力をいただいた皆さま、並びに調査にご協力をいただいた皆さまに厚く御礼申し上げます。

なお、本報告書に関して、開示すべき利益相反関連事項はありません。

## 引用文献

- Anthony, L & Chujo, K. Web Para News. <https://www.antlabsolutions.com/webparanews/> (2025年7月25日閲覧)
- Hasegawa, Y. (2017). Analyzing Explicit and Implicit Reading Questions in a Term-Exam: A Case Study. *JLTA Journal*, 20, 57-75.
- 今井裕之. (2023). 外国語科における三観点評価の課題. *KELES ジャーナル*, 8, 82-86.
- 国立教育政策研究所教育課程研究センター (2020). 『「指導と評価の一体化」のための学習評価に関する参考資料 高等学校外国語』国立教育政策研究所
- 国立教育政策研究所教育課程研究センター (2020). 『「指導と評価の一体化」のための学習評価に関する参考資料 中学校外国語』国立教育政策研究所
- 政所里佳. (2018). 「パラフレーズ・レベルによる錯乱肢の作成—語答分析に基づくテキスト理解度の診断に向けて—」. *EIKEN BULLETIN*, 30, 85-108.
- 水本篤. (2024). AI のある英語教育・研究 Let There Be AI!. *KELES ジャーナル*, 9, 52-58.
- Mizumoto, A. (2025). Langtest. <https://langtest.jp/> (2025年7月22日閲覧)
- 文部科学省 初等中等教育局. (2024.12). 「生成AIの利活用に関するガイドラインver.2.0」. 文部科学省.
- 文部科学省. (2018). 「中学校学習指導要領(平成29年告示)」. 文部科学省.
- 日本英語検定協会. (2024). 「英検2級一次試験過去問」. [https://www.eiken.or.jp/eiken/exam/grade\\_2/](https://www.eiken.or.jp/eiken/exam/grade_2/) (2025年7月22日閲覧)
- 日本リサーチセンター. (2025.7.7). 「【NRC デイリートラッキング】生成AIの利用経験 2025年6月調査」. <https://www.nrc.co.jp/report/250707.html> (2025年8月28日閲覧)
- 根岸雅史. (2017). 「テストが導く英語教育改革 「無責任なテスト」への処方箋」. 三省堂.
- OpenAI. (2024). ChatGPT- 4o (2024.5.31) [Large language model]. <https://ChatGPT.com/>
- Oxford University Press. (2020.3.10). *Oxford Advanced Learner's Dictionary 10th edition*
- R Core Team (2025). *\_R: A Language and Environment for Statistical Computing\_*. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.
- Sato, T. & TeachLex Scope Team. (2025). TeachLex. <https://teachlex.pythonanywhere.com/>
- 清水真紀. (2005). 「リーディングテストにおける質問タイプ・パラフレーズ・推論・テーマ質問と処理レベルの観点から—」. *EIKEN BULLETIN*, 17, 48-62.
- Shin, D., & Lee, J. H. (2023). Can ChatGPT make reading comprehension testing items on par with human experts? *Language Learning & Technology*, 27(3), 27-40.
- 竹内理・水本篤. (2023). 「外国語教育研究ハンドブック—研究手法のより良い理解のために—」. 松柏社.
- Terao, T., & Ishii, H. (2020). A comparison of distractor selection among proficiency levels in reading tests: A focus on summarization processes in Japanese EFL learners. *Sage Open*, 10 (1), 2158244020902087.
- 卯城祐司. (2011). 「英語で英語を読む授業」. 研究社.
- Ushiro, Y, Morimoto, Y., Hijikata, Y., Nakagawa, C., Watanabe, F., Kai, A., Shimizu, H., Koga, T., Ohno, M., Umehara, C., Hamano, Y., Narumi, T., Tao, N., Shimada, S., Kiyama, M., Suzuki, R., Kurisu, Y., & Gomi, N. (2007). What Makes Distractors Plausible in Multiple-choice Reading Tests? *Japanese Language Testing Association*, 10, 56-67. [https://doi.org/10.20622/jltaj.10.0\\_56](https://doi.org/10.20622/jltaj.10.0_56)
- Ushiro, Y., Nakagawa, C., Morimoto, Y., Hijikata, Y., Watanabe, F., & Kai, A. (2008). Effects of question types on item difficulty in two reading test formats: Open-ended and multiple-choice. *ARELE: Annual Review of English Language Education in Japan*, 19, 201-210.
- 若林俊輔・根岸雅史. (1993). 『無責任なテストが「落ちこぼれ」を作る』. 大修館書店.
- Watanabe, T., & Saida, C. (2024). ChatGPT' s Performance on the English Reading Sections of the Common Test for University Admissions in Japan. *KATE Journal*, 38, 155-168.

## 資料1: オリジナルとパラフレーズの正答選択肢一覧

| 項目番号 |      | 選択肢                                                                                              | 総語数 |
|------|------|--------------------------------------------------------------------------------------------------|-----|
| A    | (1)  | オリジナル will take part in an event at a university in Petersburg.                                  | 11  |
|      |      | パラフレーズ will join an event at a university in Petersburg.                                         | 9   |
|      | (2)  | オリジナル Hand in some forms at least seven days before then.                                        | 10  |
|      |      | パラフレーズ Make sure to give the forms in at least seven days ahead of time.                         | 14  |
|      | (3)  | オリジナル Work with two other employees to prepare for February 24.                                  | 10  |
|      |      | パラフレーズ Team up with two other employees to get ready for February 24.                            | 12  |
| B    | (4)  | オリジナル Keep an eye on how words are used in the English language.                                 | 12  |
|      |      | パラフレーズ Notice how words are used when speaking or writing in English.                            | 11  |
|      | (5)  | オリジナル a large number of people are using the word in the same way.                               | 13  |
|      |      | パラフレーズ a lot of people are using the word in a similar way.                                      | 12  |
|      | (6)  | オリジナル has gained a different meaning from its original one.                                      | 9   |
|      |      | パラフレーズ has changed from what it was originally.                                                  | 7   |
|      | (7)  | オリジナル To make the final decision about changes to the content of the dictionary.                 | 13  |
|      |      | パラフレーズ To make the last decision about updating the dictionary content.                          | 10  |
| C    | (8)  | オリジナル She received Nobel Prizes in more than one subject area.                                   | 10  |
|      |      | パラフレーズ She won Nobel Prizes in different subject areas.                                          | 8   |
|      | (9)  | オリジナル the university did not allow women to become students.                                     | 9   |
|      |      | パラフレーズ women were not allowed to be students at the university.                                  | 10  |
|      | (10) | オリジナル The food she ate was not good enough for her to stay healthy.                              | 13  |
|      |      | パラフレーズ She did not eat food that was healthy enough for her.                                     | 11  |
|      | (11) | オリジナル worked at the same university where Marie Curie was doing research.                        | 11  |
|      |      | パラフレーズ worked at the same university where Marie Curie was conducting her research.              | 12  |
|      | (12) | オリジナル Marie Curie and her sister agreed to help each other to study at university.               | 14  |
|      |      | パラフレーズ Marie Curie and her sister promised to assist each other in their university education.   | 14  |
| D    | (13) | オリジナル made some recommendations to him and his friend.                                           | 8   |
|      |      | パラフレーズ gave them a few recommendations to consider.                                              | 7   |
|      | (14) | オリジナル The way that it looked was very appealing                                                  | 8   |
|      |      | パラフレーズ Its appearance was very attractive.                                                       | 5   |
|      | (15) | オリジナル to talk with her about working at her restaurant.                                          | 9   |
|      |      | パラフレーズ to have a conversation with her about a job at her restaurant.                            | 12  |
| E    | (16) | オリジナル They left their farms and started buying things that were made in factories.               | 13  |
|      |      | パラフレーズ They moved away from their farms and began purchasing factory-made goods.                 | 11  |
|      | (17) | オリジナル The lives of people would improve if they returned to making things by hand.               | 14  |
|      |      | パラフレーズ If people returned to handmade production, their lives would become better.               | 11  |
|      | (18) | オリジナル did not have enough involvement in the production process.                                 | 9   |
|      |      | パラフレーズ lacked sufficient involvement in how things were produced.                                | 8   |
|      | (19) | オリジナル the products were expensive and not easily available.                                      | 8   |
|      |      | パラフレーズ the goods were expensive and difficult to access.                                         | 8   |
|      | (20) | オリジナル Technology developed in the nineteenth century had a significant effect on UK society.     | 13  |
|      |      | パラフレーズ The UK society was strongly affected by the technology created in the nineteenth century. | 14  |
| F    | (21) | オリジナル To congratulate her on her first year as a member of the fitness center.                   | 14  |
|      |      | パラフレーズ To give her best wishes for her first anniversary at the fitness center.                  | 13  |
|      | (22) | オリジナル arrange to be in a feature in an online magazine.                                          | 10  |
|      |      | パラフレーズ plan to be included in an article in an online magazine.                                  | 11  |
|      | (23) | オリジナル people who are introduced by other members.                                                | 7   |
|      |      | パラフレーズ people who are brought in by existing members.                                            | 8   |

(資料1 続き).....

| 項目番号 |      | 選択肢                                                                              | 総語数 |
|------|------|----------------------------------------------------------------------------------|-----|
| G    | (24) | オリジナル The location where they live is only in a certain area of Africa.          | 13  |
|      |      | パラフレーズ Their home is found only in a certain region of Africa.                   | 11  |
|      | (25) | オリジナル bonobos were seen to help other bonobos without getting a reward.          | 10  |
|      |      | パラフレーズ bonobos helped each other without expecting anything in return.           | 9   |
|      | (26) | オリジナル bonobos could feel others' feelings even more than humans.                 | 9   |
|      |      | パラフレーズ bonobos were able to understand others' emotions even better than humans. | 11  |
|      | (27) | オリジナル tend to be less friendly because there is a strong leader among them.      | 13  |
|      |      | パラフレーズ tend to show less friendliness because a dominant leader is among them.   | 12  |
|      | (28) | オリジナル Bonobos live peacefully in groups that have no strong leaders.             | 10  |
|      |      | パラフレーズ Bonobos stay in harmony in groups where no strong leader is present.      | 12  |

(注)文頭が小文字で始まっている選択肢は問題文の続きを選ぶ形の問題のため、原文のまま載せている。

## 第37回 研究助成

## B 実践部門・報告 I・英語能力向上をめざす教育実践

# 高等学校における CLIL を用いた思考力・表現力を養う授業の開発と評価

研究者: 出尾 美由紀 英国／ウォーリック大学大学院 在籍(申請時) 大阪府／港南造形高等学校 教諭

《研究助言者: 竹内 理》

## 概要

本実践研究では、高等学校において、内容言語統合型学習(Content and Language Integrated Learning: 以下 CLIL)の枠組みと、対話型鑑賞(Visual Thinking Strategies: 以下 VTS)を取り入れた英語授業を開発・実践し、生徒の思考力と表現力の向上にどのように寄与するかその効果を検証した。特に、英語が苦手な生徒でも取り組みやすいように、段階的な思考的・言語的スキュアフォームディング(足場かけ)を授業に組み込み、英語による絵画鑑賞タスクを実施した。事前・事後タスクにおける発話を質的に比較分析した結果、理由や根拠を伴う説明の増加、創造的・推測的発話の活発化、生徒同士の協働的な対話や構文的多様性の拡大が見られた。これらの変容は、語彙や文法の習得にとどまらず、英語による論理的かつ相互的な対話能力の向上を示しており、開発したVTS活用授業が高校生の思考力・表現力育成に有効である可能性を示唆している。

## 1 はじめに

大阪府立港南造形高等学校(以下、前勤務校)は、日本最大規模の美術専門高等学校であり、全国的にも高い知名度を有している。生徒たちは造形美術に強い関心と情熱を持って学んでいる一方で、英語学習への意欲は必ずしも高くないという現状がある。その背景として、多くの生徒が入学時から英語を苦手と感じており、さらに美術系大学の多くで英語が必須科目でないため、受験科目としての必要性も実感しにくいことが挙げられる。しかしながら、美術やデザインの分野そのものは国際的な広がりを持つ領域であり、作品や制作意図を自らの言葉で発信できることは、将来の大きな強みとなり得る。

このような状況において、前勤務校の英語教育では、生徒の専門分野への関心を生かしつつ、英語を「実際に使うことのできる言語」として体験できる授業設計が求められている。その一つの方法として、CLILの枠組みを取り入れてきた。具体的には、自身の作品や制作意図を英語で発表する活動や、生徒自身の専門領域の紹介ポスターを海外の生徒に向けて作成する活動などを通じて、英語を用いながら美術分野の学びを深める授業を実践してきた。

さらに、他者の作品に視野を広げ、多角的に考える力を養う活動も重要であると考え、VTSに着目した。VTSは作品を対話的に鑑賞し、他者との意見交換を通じて思考を深めていく手法であり、生徒に協働的な対話や多角的な視点を促すことができる。この取り組みは、学習指導要領で重視されている「思考力・判断力・表現力」のうち、特に「思考力」と「表現力」との親和性が高く、それらの向上に資することが期待される。しかしながら、生徒の英語習熟度に応じて、段階的に言語的支援を取り入れながらVTSを実践する工夫が必要である。本研究では CLIL の枠組みを用いて VTS を取り入れた授業を開発・実践し、生徒

の思考力と表現力がどのように変容するかを評価・検証することを目的とする。

## 2 研究の背景

### 2.1 CLILの理論的枠組み

CLIL は、教科内容を外国語で教える教育アプローチであり、内容の学習と言語の学習という二重の焦点を持ち(Coyle et al., 2010)、内容と言語は同等に重要であるとされている(Cook, 2016)。その理論的基盤となるのが「4Cs フレームワーク」である。Coyle et al.(2010, p. 41)によれば、この枠組みは Content, Communication, Cognition, Culture(または Community)の四つの要素の相互作用を通じて学びを深めることを目的としている。Contentは教科の学習内容を指し、Communicationはその内容を理解・表現するための言語学習・言語使用を意味する。Coyle et al.(2010)は、「両者は互いに支え合い、言語を使うことで内容理解が深まり、内容学習を通じて言語力も向上する」と指摘している。Cognitionは、学習と思考に関わる幅広いプロセスを包含し、Bloomの思考の分類に基づく低次思考力(Lower-order thinking skills: 以下LOTS)の記憶・理解・応用から高次思考力(Higher-order thinking skills: 以下HOTS)の分析・評価・創造に至るまでの多様な思考スキルを含むものであり、これらすべてが効果的な学習に不可欠な要素である(Coyle et al., 2010)。Culture/Communityは、異文化理解やグローバルシティズンシップの育成をめざす観点であり、教室内の協働的な学びと同時に、学習が社会や世界とどのように結びつくかという広がりを意識させる枠組みとして位置付けられている(池田, 2011)。これらの4Csの要素が意識的に授業計画や教材、活動に統合されていることが、非常に重要であり、CLILを他の教育法と区別する際立った特徴である(Coyle et al., 2010; Ikeda, 2021)。

しかしながら、理論上は4Csすべてが均等に重視されるべきであるが、実際の授業ではContentとCommunicationのどちらか一方が主目的となることが多い(Ikeda, 2021)。教科内容が主目的となる内容中心型はHard CLIL、言語能力の習得が明確な目標で、内容は言語学習の手段として用いられる言語中心型はSoft CLILと呼ばれる(池田, 2011)。特にSoft CLILについてIkeda(2021, p.16)は次のように定義している：

Soft CLILとは、内容に大きく依存する言語教授法であり、学習者は目標言語を使って仲間と活発に言語で交流し、認知的に高度な活動や文化的に豊かな活動に参加しながら、言語能力、教科に関する理解、そして他の場面にも応用できる多目的なスキルを同時に伸ばしていくものである。(筆者訳)

本研究で開発・実践する授業は、様々な国で描かれた絵画を対話的に鑑賞しながら、生徒が観察・推論・解釈を行い、それを英語で表現することを主な目的としている。このような活動は、まさにIkeda(2021)が述べるSoft CLILの定義に合致しており、本研究の実践はSoft CLILに位置付けられる。次節では、その枠組みを授業で具体化する手法として導入したVTSの概要とその教育的意義を示す。

### 2.2 VTSの概要と教育的意義

VTSは、Abigail Housen と Philip Yenawine によってニューヨーク近代美術館の教育プログラムの一環として開発された教育手法である(Visual Thinking Strategies, n.d.)。VTSは絵画や写真などの視覚素材を用い、生徒が自由に観察し対話的に発言し合うことで深い理解を促すものである。教師は知識の一方的な伝達者ではなく、ファシリテーターとして①What's going on in this picture? ②What do you see that makes you say that? ③What more can we find?という三つの質問を繰り返し投げかける。これにより、

生徒は観察の根拠を言語化し、他者の意見を取り入れながら思考を発展させることができる教育手法である (Yenawine, 2013)。このような活動は、美術専攻である前勤務校の生徒の興味・関心と非常に合致しており、英語が苦手な生徒にとっても、学習意欲の向上に寄与することが期待される。さらに、対話的に鑑賞を行うスキルは、美術の授業や日本語での作品鑑賞などにも役立つことが予想され、結果として多教科にまたがる学習能力の向上にも貢献し得る。これらの理由から、VTS を英語授業に導入することを決定した。

先行研究によれば、VTS は生徒の観察力、批判的思考力、表現力の向上に寄与することが報告されており (Yenawine, 2013)、また美術教育以外の分野でもコミュニケーション能力や観察力を高める効果が示されている (Aspden et al., 2022; Poirier et al., 2020)。このように、VTS が促進する対話的な思考プロセスを通して様々なスキルを伸ばすことが期待され、英語授業に導入することで、CLIL における 4Cs のすべてと親和性を有すると考えられる。すなわち、絵画に関する議論を通じて内容 (Content) に迫り、他者の視点を尊重し合いながら協働的に意味を構築する協働学習活動 (Community) を通して、観察・推論・根拠・提示という思考のプロセス (Cognition) を深め、それらを言語化してやり取りすることで言語使用 (Communication) を活性化させることができる。

その中でも特に、Cognition と Communication との親和性が高い点が注目される。VTS の中心となる質問① What's going on in this picture?, 質問③ What more can we find? は、生徒の観察や創造的・推測的な思考の表出を促し、質問② What do you see that makes you say that? ではその理由や根拠の提示を求めることで、Cognition の発展に大きく寄与すると考えられる。また、それらの思考を対話を通して英語で言語化させることにより、Communication の発展にもつながる。VTS によって育成される Cognition と Communication は、学習指導要領 (文部科学省, 2018) で重視されている「思考力・判断力・表現力」のうち、特に「思考力」と「表現力」に呼応していると考えられ、VTS を活用した英語授業によって、これらの向上が期待できる。さらに、学習指導要領 (文部科学省, 2018) が示す知識の習得にとどまらず、「主体的・対話的で深い学び」の実現においても大きく寄与することが期待される。VTS はまさにそのような学習観を体現しており、日本の英語教育がめざす学びの姿に沿った実践であると言える。

したがって、VTS は CLIL の 4Cs フレームワークと呼応すると同時に、日本の英語教育が求める「思考力・表現力の向上」や「主体的・対話的で深い学び」を具体化する一手法になり得る。しかしながら、特に前勤務校のように英語を苦手と感じている生徒が多い教育現場にとっては、十分に思考を促す課題を提示しつつ、同時に適切な言語的支援を行う必要がある。Coyle et al. (2010, p. 42) も指摘するように、学習者が認知的に挑戦されつつも、言語面で適切に支援される環境を戦略的かつ体系的に整えることが、英語の授業で VTS を成立させるために必須であると言える。

## 2.3 本研究の位置付けと研究目的

以上を踏まえ、本実践研究では高等学校の英語授業において、CLIL の枠組みを基盤とし、英語が苦手な生徒でも取り組みやすいよう、段階的な思考的・言語的スキュアフォールディングを組み込んだ VTS 活用授業の開発・実践を行う。そして、開発した授業が生徒の思考力と表現力にどのような変容をもたらすかを分析することで、授業の評価・検証を行うことを目的とする。本研究のリサーチクエスション (RQ) は以下の通りである。

**RQ** 開発した VTS 活用授業は、生徒の思考力・表現力の発展にどのような影響を与えうるか。

また、授業設計が生徒に与える影響を探る手掛かりとして、以下の仮説を設定した。前節で述べた通り、VTS との親和性の高さと、目的である「思考力」と「表現力」にも呼応しているという点から、4Cs の中でも Cognition と Communication の 2 点に焦点を置いた。思考力 (Cognition) では、生徒の発話における理由や根拠の提示、創造的・推測的思考の表出など思考力の発展を、そして表現力 (Communication) では、

賛同表現の多様化や異なる意見の提示を通じた対話構築能力、発話の文構造や内容表現の明確化など表現力の発展を分析する。具体的には、以下の変容が見られる場合、授業設計が意図した学習効果に寄与している可能性があると考ええる。

思考力(Cognition)に関する仮説:

- 仮説1:生徒の発話に、理由や根拠を明示する内容が増加する。
- 仮説2:生徒の発話に、創造的・推測的な思考が含まれる内容が増加する。

表現力(Communication)に関する仮説:

- 仮説3:生徒は賛同表現の多様化や異なる意見の提示を通して、対話をより豊かに表現できるようになる。
- 仮説4:生徒は発話を文として組み立て、内容をより明確に表現できるようになる。

## 3 研究方法

### 3.1 参加者

参加者は、前勤務校3年生のうち「実用英語」を履修している9名である。生徒は学習活動の一環として、VTSを取り入れた授業および事前・事後タスクに参加した。授業実践者である筆者は2023年度から生徒の英語指導に当たっており、これまでの指導経験から、活動の中心である「話すこと」「聞くこと」における生徒の英語力は、ほとんどの生徒が、事前に準備すれば日常的な事柄について基本的な表現を用いて複数文で表現でき、教師の指示やゆっくりと発話される内容であれば身の回りのことを聞き取ることができる(CEFR A1, CEFR-J A1.3レベル; 大阪版CAN-DOリスト参照)。一方で、適切な表現や文構造を用いて即興で説明したり、対話を発展させたりすることはまだ困難である。

事前・事後タスクの録画データ取り扱いについては、文字起こしデータのみを仮名化して使用し、本研究以外の目的には使用しないことを口頭にて説明し、参加生徒全員から同意を得た。また、学校長からもデータ使用の許可を文書で取得した。

### 3.2 事前・事後タスク

事前・事後タスクでは、質問①～③を生徒自身が活用し、ペアで互いに鑑賞を促進し合いながら観察や思考を深めるVTS活動を行った。基本的なVTSでは、教師がファシリテーターとなり生徒の発話を促すが、これまでの指導経験から、教師がクラス全体に質問を投げかけてしまうと、クラス全体の前で自身の意見を安心して発表することが難しく感じる生徒が多く、決まった生徒のみが意見を述べる可能性があると考えられた。そこで、ペア活動にすることで、生徒がより安心して発話できるとともに、発話のターンテイキングも平等に行うことが可能となる。さらに、質問を生徒自身が能動的に使用することで、より対話的なやり取りが期待される。以上の理由から、ペアでVTSを用いた鑑賞活動を実施することに決定した。

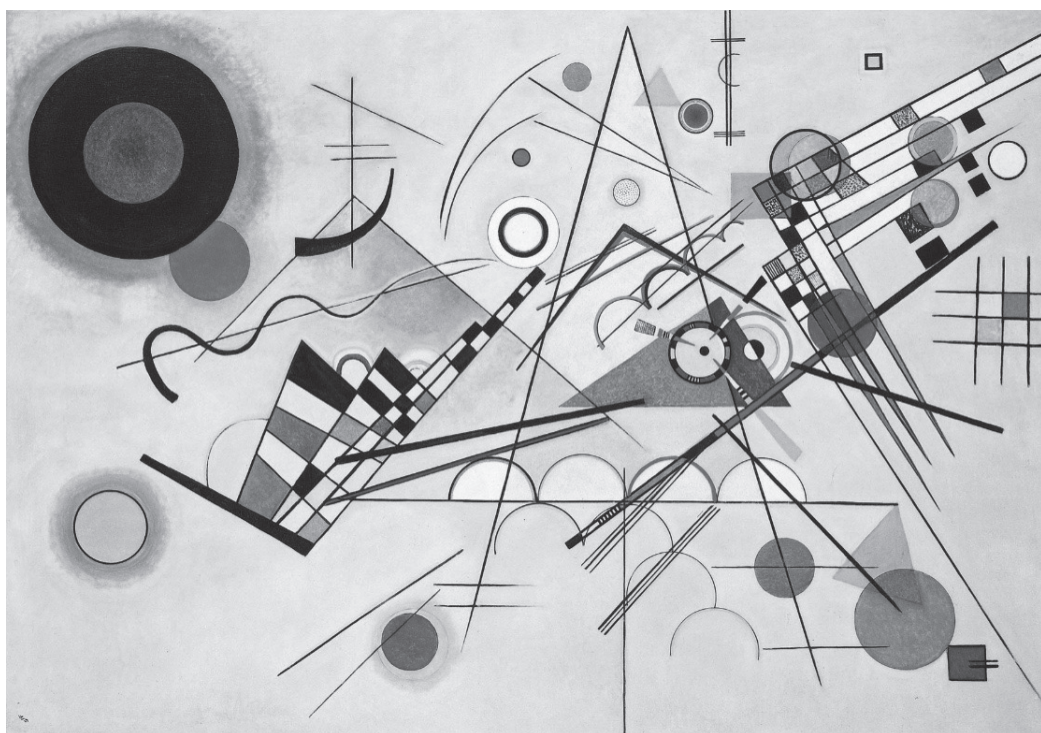
使用した絵画は、A『こんにちは、クールベさん』(ギュスターヴ・クールベ作、人物・具象画)、B『花嫁』(マルク・シャガール作、人物・抽象画)、C『コンポジションVIII』(ワシリー・カンディンスキー作、抽象画)の3作品とした(図1～3参照)。生徒を2人1組のペアとし、絵画A～Cを用いたVTS活動を行った。鑑賞時間は各作品につき2分間、合計約6分間とした。ペアはランダムに決定し、事後タスクでは作品の提示順序もペアごとに無作為に変更した。質問①～③と日本語訳を印刷したカード(図4参照)を生徒に配布し、必要に応じて参照できる形とした。作品はパソコン画面に提示し、そのパソコンを用いて活動の様子を録画した。



■図1: ギュスターヴ・クールベ『こんにちは, クールベさん (Bonjour Monsieur Courbet)』  
(1854年, 油彩, ファーブル美術館蔵) 出典:Wikimedia Commons



■図2: マルク・シャガール『花嫁 (La Mariée)』(1912年, 油彩, ポンピドゥー・センター蔵)  
出典: ポンピドゥー・センター (<https://www.marcchagall.net/la-mariee.jsp> より取得)



■ 図3: ワシリー・カンディンスキー『コンポジションVIII (Composition VIII)』(1923年, 油彩, 個人蔵)  
出典: Wikimedia Commons

- What's going on in this picture?  
この作品の中で何が起こっていますか？
- What do you see that makes you say that?  
なにを見てそう思いましたか？
- What more can we find?  
もっと発見はありますか？

■ 図4: 質問①～③と日本語訳を印刷したカード

### 3.3 授業内容

本実践は1学期間にわたり実施した。授業は週に1回、2時間連続(1コマ50分×2)で行われ、そのうち約1～1.5コマをVTS活用授業に充てた。主な教材として、国立国際美術館から借用した「アートカード」65枚を用い、様々な活動を通じてVTSの実践を行った。

授業全体の設計は、VTSの基本原則である観察や推論の根拠を言語化し、他者の意見を取り入れながら思考を発展させる(Yenawine, 2013)ことを、段階的に習得できるように構成した。まず第1段階では、絵画そのものを客観的に描写する練習を行い、第2段階では、鑑賞中に抱いた感情を表現するとともに、その根拠を絵画の描写に基づいて説明することを支援した。さらに、自分の意見や感情を述べるだけでなく、相手の意見に賛同したり、受け入れつつ異なる視点を提示したりする活動を取り入れた。最終の第3段階では、絵画から想像できる物語や情景を表現する活動を行った。

各回の授業計画は池田(2016)のCLIL授業設計フレームワークを活用し、開発を行った。生徒の英語力を鑑みて、必要な言語要素は明示的に指導し、それらを実際に使用できる様々なVTS活動を設計・実践した。1学期における授業計画の概要を表1に示し、具体的な授業内容はCLILの授業設計のフレームワーク(池田, 2016)とあわせて以下に記す。

■表1: 授業計画の概要

| 段階   | 週  | VTS               | 言語要素                                                         |
|------|----|-------------------|--------------------------------------------------------------|
| 導入   | 1  | 対話型鑑賞導入           | —                                                            |
|      | 2  | 事前タスク             | —                                                            |
| 第1段階 | 3  | 絵画の人物やものの描写       | 状態: There is a woman.                                        |
|      | 4  |                   | 行動: The woman is sitting on a chair.                         |
|      | 5  | 絵画の色や雰囲気の描写       | 色: This uses a lot of blue.                                  |
|      | 6  |                   | 雰囲気: This looks bright.                                      |
| 第2段階 | 7  | 気持ちの描写            | This makes me feel _ because...<br>(状態・行動・色・雰囲気などを理由として述べる。) |
|      | 8  | 相手の意見に賛成／異なる意見を提示 | I agree with you. / I see what you mean. In my opinion, ...  |
| 第3段階 | 9  | 想像していることの描写       | I think..., I imagine...                                     |
|      | 10 |                   |                                                              |
| まとめ  | 11 | 事後タスク             | —                                                            |

#### 3.3.1 VTSの導入

[第1週目] VTSの概要を説明した。生徒の日本語での鑑賞経験を引き出ししながら、多様な解釈が存在してよいこと、対話的に鑑賞を行うことの重要性を説明した。また、VTSを促す質問①～③の意味の確認と発音・使い方の練習を行い、実際の絵画を用いてクラスでVTS活動を行った。

[第2週目] 事前タスクを実施した(3.2参照)。事前タスクは、生徒の思考力・表現力の変化を検証するためだけでなく、生徒が不足している語彙や表現を明らかにする診断的役割も持つ。その結果に基づき、以下の授業内容を設計・実践した。

#### 3.3.2 【第1段階】絵画の描写活動

[第3・4週目] 絵画そのものを客観的に描写する練習を行うため、人物の行動や物の状態を描写するのに

必要な言語要素を復習し、それを活用したアートカードを使った推測ゲーム形式の活動を行った。まず一人の生徒が手元のカードから1枚を選び、心の中で決める。その後、そのカードに描かれた内容を、「動作」や「状態」に関する三つの英文で説明する。ただし、この段階では色に関する表現は使用しないように指示した。説明を聞いた他の生徒は、提示された情報を手掛かりにしてどのカードであるかを推測し、合図に合わせて一齐に指をさす。正しいカードを当てられた場合、当てた生徒はそのカードを獲得できる仕組みとした。

〔第5・6週目〕さらに描写活動を深めるため、作品の色や雰囲気(bright, dark 等)を表す語彙や表現を提示し、前回の推測ゲームを発展させた活動を行った。生徒は1枚の絵を決め、その絵を三つの英文で描写する。この際、前回は動作と状態のみを用いたが、今回はそれに加えて、目標語彙である色や雰囲気も含めて表現することを課題とした。色の描写をすると、その文章のみの描写でカードがわかってしまう可能性があるため、生徒が描写をしている間は、他のメンバーは机に伏せ、3文を聞き終わってから、合図とともに顔を上げ、絵を探し出し、該当する絵を取り上げることとした。前回と同様に正しい絵を当てることができれば、その絵を獲得できる仕組みとした。

■表2: 【第1段階】CLIL授業設計

| Content                                                       | Communication                                                                                                                                                    | Cognition                                  | Community Culture                               |
|---------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|-------------------------------------------------|
| <b>【宣言的知識】</b><br>絵画にどのような人物や物が描かれているのかを知る。                   | <b>【言語知識】</b><br>〔第3・4週目〕<br>語彙：人物や物、基本的な動詞<br>表現：There is .... S is doing ....<br>〔第5・6週目〕<br>語彙：色、雰囲気を表す形容詞<br>表現：This uses (a lot of) ....<br>This looks .... | <b>【LOTS】</b><br>理解：絵画の描かれている人物や物の特徴を描写する。 | <b>【協働学習】</b><br>ペア<br>グループ                     |
| <b>【手続的知識】</b><br>絵画の人物や物について、自分と他者の見方を照らし合わせ、絵画の多様な捉え方を理解する。 | <b>【言語技能】</b><br>話すこと<br>聞くこと                                                                                                                                    | <b>【HOTS】</b><br>評価：自分なりの絵画の解釈を説明する。       | <b>【国際理解】</b><br>世界各地の多様な作品を通じて、国際的な芸術の多様性に触れる。 |

### 3.3.3 【第2段階】感情の表現活動

〔第7週目〕この週より、鑑賞中に抱いた感情を表現することに焦点を置き、その感情の根拠を絵画の描写に基づいて説明できるよう支援した。これは特にVTSの質問②を意識した内容である。生徒の気持ちや感情を表現する語彙(scared, peaceful, lonely, depressed, amused, confused, curious)を目標語彙として設定し、これらの単語が想起される絵画作品を用いて各単語の意味やニュアンスを確認した。次に、絵画を観察しながら、“This makes me feel...because...”の表現を使って、自分の言葉で作品がどのような感情や気持ちを喚起するかを表現する活動に取り組んだ。このとき、これまでに学習した語彙や表現を活用し、絵画の描写を感情の根拠として使うよう促した。

〔第8週目〕上記の活動に加え、相手の意見に対して賛同や異なる意見を示す表現を導入し、自分の感じたことや考えを対話的に述べる活動を実施した。異なる意見や感想を表す表現については、絵画の鑑賞で感じる感情が人それぞれ異なることを尊重し、否定するのではなく、相手の意見を受け入れつつ自分の意見も伝える姿勢を重視した。

■表3: 【第2段階】CLIL授業設計

| Content                                              | Communication                                                                                                                                                                                                                        | Cognition                                      | Community Culture                               |
|------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------|-------------------------------------------------|
| <b>【宣言的知識】</b><br>感情表現に根拠となる絵画の要素（人物や物、色、構図など）を理解する。 | <b>【言語知識】</b><br>[第7週目]<br>語彙：感情を表す形容詞<br>表現：This makes me feel ... because...<br>[第8週目]<br>表現：<br>（賛同）I (completely / kind of) agree with you.<br>（異なる意見）I like your idea. But to me, ... / I see what you mean. In my opinion, ... | <b>【LOTS】</b><br>理解：絵画の描かれている人物や物の特徴を描写する。     | <b>【協働学習】</b><br>ペア<br>グループ                     |
| <b>【手続的知識】</b><br>他者の感情表現を聞き、異なる見方や感情の捉え方を理解する。      | <b>【言語技能】</b><br>話すこと<br>聞くこと                                                                                                                                                                                                        | <b>【HOTS】</b><br>分析・評価：絵画の描写から自分の感情の根拠を説明している。 | <b>【国際理解】</b><br>世界各地の多様な作品を通じて、国際的な芸術の多様性に触れる。 |

### 3.3.4 【第3段階】創造的な表現活動

[第9・10週目] アートカードを用いて絵画の場面や背景について、生徒自身の想像力を働かせて説明する活動を行った。これはVTSの質問①③において、描写にとどまらず、より創造的な発話を促進することをねらいとした。また、相手の意見に対しては、第2段階で習得した賛否を表すフレーズを活用し、互いの考えを尊重しながら対話を進めることを促した。最後に、ペアでまとめた内容を基に4コマのストーリーを作成し、クラス全体に発表した。

■表4: 【第3段階】CLIL授業設計

| Content                                                 | Communication                                                         | Cognition                                                   | Community Culture                               |
|---------------------------------------------------------|-----------------------------------------------------------------------|-------------------------------------------------------------|-------------------------------------------------|
| <b>【宣言的知識】</b><br>絵画の場面や背景に関する具体的な情報や特徴を知る。             | <b>【言語知識】</b><br>[第9・10週目]<br>表現：I think... / Maybe... / I imagine... | <b>【LOTS】</b><br>理解：絵画の描かれている人物や物の特徴を描写する。                  | <b>【協働学習】</b><br>ペア<br>グループ                     |
| <b>【手続的知識】</b><br>絵画の特徴を基に想像を広げ、他者とのやり取りを通して考えを発展・共有する。 | <b>【言語技能】</b><br>話すこと<br>聞くこと                                         | <b>【HOTS】</b><br>創造：観察に基づき、絵画の背景や描かれていない部分について想像・創造することができ。 | <b>【国際理解】</b><br>世界各地の多様な作品を通じて、国際的な芸術の多様性に触れる。 |

### 3.3.5 まとめ

[第11週目] 事後タスクを実施し(3.2参照)、成績の一部として反映するため、実践的で簡潔な評価基準を用い、評価を行った。

### 3.4 評価・検証方法

本研究では、仮説1~4(2.3参照)に基づき、CognitionとCommunicationの観点から事前・事後タスクを比較し、生徒の変容を分析することで、開発したVTS活用授業が生徒の思考力・表現力の発展にどのように影響しうるかを検証する。

分析手法としては、録画の文字起こしデータを基に、思考力・表現力の変化を仮説1~4の観点から質的に談話分析を行う。生徒の思考力・表現力の変容を分析するにあたっては、生徒の思考過程や発話内容の質的特徴、対話における柔軟性や創造性の変容を詳細に捉える必要があるため、単に発話回数や文量などの数値に還元する量的分析では把握が難しいと判断し、質的アプローチを採用した。

## 4 結果および考察

以下では、事前・事後タスクの比較・分析で得られた結果を仮説ごとに示し、それぞれの結果について考察する。分析では、生徒の発話の例もあわせて提示し、具体的な変化の様子を示す。なお、生徒の名前はすべて仮名化しており、抜粋内ではイニシャルを用いて記載している。

### 4.1 理由や根拠を含む発話の変容(仮説1)

VTSで用いられる対話促進のための三つの質問のうち、質問②What do you see that makes you say that?は、発話の根拠を視覚的証拠に結びつけることを促し、分析的観察を支える質問である。事前タスクの結果からは、生徒が理由や根拠を含めて発話する際、多くの場合(5ペア中4ペア)がこの質問②をきっかけとして根拠提示を行っていた。例えば、絵画Bに関する次のやり取りでは、Kokiがこの絵はパーティで夜であると述べた後、Ayaが質問を挟むことで初めて根拠が引き出されている(下線部参照)。

#### 例1: 絵画Bの理由提示(事前タスク)

1. K: Hold violin [laugh] play the vi- violin and other people dancing
2. A: Ahh dancing dancing
3. K: It's party [laugh]
4. A: うん party party [laugh] uuum fish fish is dancing
5. K: Fish is dancing [laugh]
6. A: [laugh] errm black
7. K: Black
8. A: Black
9. K: Night
10. A: Night
11. K: It's night
12. A: It's night. What do you see that makes you say that?
13. K: Because it's um.. uum
14. A: [nodding]
15. K: Errm.. many animals and people
16. A: うん
17. K: And it's um gor- the woman is gorgeous
18. A: うん
19. K: Dress
20. A: Umhum

このやり取りでは、質問がきっかけとなり、Kokiは具体的な視覚情報(動物や人物の多さ、女性のドレスの特徴)を根拠として提示している。すなわち、質問が発話の分析的深まりを促す触媒として機能していることが分かる。

しかしながら、このように適切に根拠が述べられた例は少数であり、大半はこの質問に対する応答が難しく、長い沈黙や不明確な回答となっていた。例えば、絵画AについてSuzuが感想を述べた後の次のやり取りがある。

#### 例2: 絵画Aの理由提示(事前タスク)

1. S: Fa- family?
2. H: Family? [laugh]
3. S: [laugh]
4. H: What do you see that makes you say that?
5. S: うーん
6. (6.0)

注. line 6 の(6.0)は、沈黙の長さを秒で示す(6.0=6秒)。

この例から、一部の生徒は、根拠を視覚的に探し出し、理由付けを行うことを難しく捉えていることが分かる。単語やジェスチャーを補助的に用いることもなく沈黙していることから、言語能力の不足というよりも、質問の意図は理解していても、瞬時に自分の意見の理由を関連付けて説明するスキルが十分に形成されていない可能性が示唆される。

VTS活用授業を通して、生徒は単なる描写ではなく、感想や想像の理由付けとして統合する活動を継続的に経験した。その結果、事後タスクでは、質問がなくても自発的に根拠を組み込む発話が多く見られるようになった。例えば、事前タスクにおいて根拠提示に苦戦していたMamiは、事後タスクでは授業で導入したThis makes me feel... because...を活用しながら、自らの感情や理由を自然に結び付けて発話している。

#### 例3: 絵画Bの理由提示(事後タスク)

1. M: I え This picture,
2. F: うん
3. M: Makes me feel.. scared
4. F: Ahh me too
5. M: こわい.. err because.. err back back [laugh]
6. F: うん
7. M: Err ghost many
8. F: Ahh

この発話例では、Mamiが自身の感情「怖い」という感想を述べた後に、becauseを用いてその理由を説明しようと積極的に試みている。言語面では発展途上の部分も見られるが、自発的に感情と視覚的特徴(幽霊の存在)を結び付けていることが分かる。このような自発的な理由付けの試みは、思考力の育成において重要な進展を示している。

さらに、以下のやり取りでは、絵画Cについて、Kokiが視覚情報を根拠として提示し、Erikaがそれを受けて補足することで、より豊かな理由付けが展開されている。

#### 例4: 絵画Cの理由提示(事後タスク)

1. K: I saw.. it's looks noi- too noisy
2. E: Ahhh yeah
3. K: Because um.. many writes and.. many writes and.. circle.. two circle [laugh]
4. E: Yeah this err I agree this pictures are circle and err
5. K: Square
6. E: Yeah square err this pictures err uses circle and square and tri-
7. K: Triangle
8. E: Triangle mmm only
9. K: Only

このやり取りでは、ErikaがKokiの理由に補足を試みることで、Kokiから追加の意見が引き出され、共同での理由付けが深化していることが分かる。言語的にはまだ発展途上で不完全な表現も見られるが、両者が視覚情報を共有しながら分析を進める中で、相互の発話が理由付けを拡張・精緻化している。

例3・4のような変化は、VTSを活用した絵画鑑賞活動が、生徒の認知過程における理由付けスキルを徐々に定着させたことを示唆している。Bloomの分類の観点から見ると、事前タスクでは発話の多くが単なる描写にとどまり、「理解」のレベルに位置していた。それに対し事後タスクでは、HOTSにあたる「分析」や「評価」へと移行し、絵画の要素、自身の感情、推論といった複数の情報を関連付けながら理由付けを行う様子が見られた。このような変化は、CLILの枠組みにおいて「内容学習(芸術鑑賞)」と「言語使用(理由付け表現)」が相互に促進され、統合的に発展する可能性を示している。

## 4.2 創造的・推測的な思考を含む発話の変容(仮説2)

さらに、理由付けスキルの向上に加えて、発話内容にも質的な変化が見られた。具体的には、単なる視覚的描写にとどまらず、観察した情報を手掛かりに想像や推測を取り入れる場面が事後タスクにおいて増加しており、より創造的な思考の発現につながっていた。

事前タスク(例5)において、特に具象画である絵画Aにおいては、質問①What's going on in this picture?に対し、以下のような人物や動物の存在などの事実描写が中心であった。

#### 例5: 絵画Aの描写(事前タスク)

1. S: うーん .. What's going on in this picture?
2. H: うーん (5.0) Three people [laugh]
3. S: Three man?
4. H: Three man three man
5. S: うん
6. H: And one dog
7. S: [laugh] うん
8. (9.0)

注. line 8 の(9.0)は、沈黙の長さを秒で示す(9.0=9秒)。

このやり取りでは、視覚的要素の指摘に終始しており、登場人物の背景や行動の理由、場面に関する推測、あるいは感情的な解釈などはほとんど見られなかった。こうした発話は、絵画を事実情報としてのみ捉えていることを示し、解釈や想像を交えた高次の認知活動にはつながっていないと言える。5ペア中2ペアは

この例と同様に、描写で発話が完結しており、残りの3ペアには登場人物の状況や行動を想像する発話が見られたものの、それらは単発的であった。つまり、事前タスクでは多くの生徒が、視覚情報を基に新たな意味や物語を構築する段階には至っておらず、創造的・推測的思考の発現は限定的であったと言える。

一方、事後タスク(例6)では、同じく人物と犬に関する描写から(line 1-3)、登場人物の感情や状況を推測し(line 4-5)、さらにその推測を基に場面を物語的に展開する様子が観察された(line 7-8)。

#### 例6: 絵画Aの推測を含む発話(事後タスク)

1. A: Three man three man
2. S: And one dog
3. A: [laugh] one dog yes
4. S: But this man is so s- sad
5. A: Sad
6. S: Mmmm
7. A: Trip?
8. S: Oh! Ohhh
9. A: And? Two two man is..
10. S: うん
11. A: [gesture] mmm difficult clothes
12. S: Ahh ahh
13. A: One people is..
14. S: One people
15. A: Trip
16. S: One people go tr- Ahh trip.. えー?
17. A: えー? ah big.. リュック
18. S: Ah! Big bag!
19. A: Big bag one is have big bag
20. S: Oh oh oh oh, oh oh [laugh]

この例では、Suzuがまず男性の感情(sad)を推測し、その発話をAyaが受け止めた上で、出来事(trip)の可能性を提示している。さらに、男性の服装や持ち物(big bag)といった具体的描写が新たに付け加えられ、それらが推測の根拠として機能している点が注目される。このように、描写と推測が相互補完的に積み重なり、発話が場面描写から簡単な物語的展開へと移行していることが分かる。生徒同士が相手の視点を取り込みながら内容を広げていくプロセスは、創造的・推測的思考が対話の中で共同構築されていることを示しており、事後タスクにおける質的な変化の一端と言える。

さらに、より創造的な発話が促されやすい抽象画である絵画Cにおいても、事前タスク(例7)では絵画の一部の特徴を基にした短い推測や見立ては見られたものの、それらは短く簡潔なものにとどまっていた。

#### 例7: 絵画Cの推測(事前タスク)

1. H: What's going on in this picture?
2. S: ん? mountain?
3. H: mountain
4. S: Mountain... sun
5. H: Ah ここ
6. S: うん
7. H: sea
8. S: [gesture]
9. H: 海

このやり取りでは、mountain, sun, seaといった見立てが次々と挙げられてはいるものの、それらがそう見える理由や、相互の関係性については全く言及されていない。また、発話は名詞や単語レベルでのやり取りが多く、文としての展開や論理的な接続が欠けているため、場面や意味の広がりにはつながっていない。つまり、事前タスクでは、抽象画という多様な解釈が可能な題材であっても、その潜在的可能性を十分に活用した推測や物語の構築には至らず、視覚的要素の単発的な列挙にとどまっていたことが分かる。

一方、事後タスク(例8)においては、推測や見立てが色彩や形状といった具体的かつ客観的な根拠と結びつく傾向が顕著に現れ、生徒同士がそれらの情報を基に相互に補足し合うやりとりが増加した。例えば、以下は絵画CにおいてNamiとHanaが対話的に推測を伴う場面である。

#### 例8: 絵画Cの創造的発話(事後タスク)

1. N: Ahh.. this looks like sun because this uses red
2. H: I agree [laugh]
3. N: [laugh]
4. H: Mm I think it is moon
5. N: Ohh
6. H: Because this colour is yellow
7. N: Oh nice ah.. ummm
8. (4.0)
9. N: [pointing] This is mountain!
10. H: Yeah [nod] I [laugh]
11. N: [laugh]
12. H: I agree
13. N: Ahhh..
14. H: I- ah I think this is.. sea
15. N: Ohhh nice
16. H: Because this is blue
17. N: Ohh I see. ohh

注. line 8 の(4.0)は、沈黙の長さを秒で示す(4.0=4秒)。

この例では、Namiが色彩(red)を根拠にsunと見立て、Hanaも色彩(yellow, blue)を根拠にmoon, seaと推測している。事前タスクと比較すると、単なる創造的な見立ての描写から一歩進み、明確な理由付けがなされた上で相手と意見交換が行われる形態へと変化している。また、仮説1で明らかになった理由付けの能力の向上を土台として、仮説2で見られた創造的・推測的発話に根拠と説得力が加わった結果であると考えられる。つまり、生徒はただ自由に想像する段階を超え、論理的根拠に基づいた発話を通じて互いの意見を補強しながら話を組み立てているのである。このため、推測がより論理的かつ対話的に展開され、単なる思いつきや感覚的な見立てを超えて、根拠に裏打ちされた協働的な意味付けが形成されていることが分かる。

例6・8のような発話は、Bloomの思考の分類でいうところの「理解」から「応用」「分析」、さらに「創造」へと段階的に認知レベルが向上していることを示唆する。描写(知識・理解)から推測・理由付け(応用・分析)を経て、物語的展開や相互補足のある創造的な対話へと発展している。したがって、本活動における創造的・推測的な発話の増加は、単なる語彙量や文法知識の向上だけでなく、芸術鑑賞という内容領域を媒介としたHOTSの発達と、言語的相互作用の深化が相互に強化し合った結果であると考えられる。

### 4.3 協働的な対話構築能力の変容(仮説3)

生徒の思考力の向上だけでなく、表現力の発展にも顕著な変化が見られた。特に事後タスクにおいては、単なる意見の表明にとどまらず、賛同表現の多様化、異なる意見の提示、自発的な質問の増加といった、

対話を積極的に深める姿勢が顕著となった。

#### 4.3.1 賛同表現の多様化

変化の一つとして、賛同表現の多様化が挙げられる。事前タスクにおいても、多くの生徒はパートナーの意見に対して何らかの反応を示していたものの、その多くは Ahhh, Ohhh といった感嘆詞や、うなずきなどの非言語的表現に限られていた。例えば、Nami は次のようにやり取りを行っている。

##### 例9: 絵画Bの観賞(事前タスク)

1. E: Painting?
2. N: Painting ahhh! [nodding]
3. E: I think it is wedding
4. N: Ah wedding! Ohhh [nodding]
5. E: Music and bouquet
6. N: Ah! Bouquet okay okay

この例では、Nami は相手の発言に対し感嘆詞で反応し、うなずきなどの非言語的行動も交えながら共感を示している。しかし、言語面では単語や短い感嘆にとどまり、意見に対する明確な同意や理解の表明は限定的である。

一方、事後タスクにおいては、これまで単に感情的な反応にとどまっていた賛同表現が大きく進化し、授業で習得した I agree, I see などの述語を伴う明確な言語表現が頻繁に用いられるようになった。これにより、相手の発言を積極的に受け入れ理解し、対話の中でその意見を支持する意思を言葉で示すことが可能となっている。Nami の事後タスクでの発言は次の通りである。

##### 例10: 絵画Cの観賞(事後タスク)

1. N: Ahh this mm? there are many circle looks like.. star
2. H: Ahh but this is.. not night
3. N: Ahh [nod] I see ahhhh I agree I agree. oh
4. H: So I think they are little water[gesture]
5. N: Ah water

この例では、Nami は相手の異なる見解を受け入れつつ、I see, I agree といった表現で理解と同意を明確に示している。また、うなずきや感嘆詞といった非言語的要素も引き続き使われており、対話がより豊かで自然なものになっていることが分かる。

このように、賛同表現は単なる感情的反応から、より討論的かつ協働的な理解の確認へと変化した。これにより、生徒同士が互いの意見を積極的に受け入れ合いながら対話を深める場面が増え、授業で学んだ言語表現が実践の場で自然に活用されていることが明らかとなった。

#### 4.3.2 異なる意見の提示

さらに、パートナーとは異なる意見を積極的に提示する発言が顕著に見られるようになった。事前タスクでは、生徒同士のやり取りは主に相手の意見への同意や感嘆を中心に進み、異論や対立的な視点を示す場面はほとんど確認されなかった。しかし、事後タスクにおいては、But I think..., But in my opinion... といったフレーズを用い、自分の考えや異なる見解を明確に示す例が複数観察された(例11・12下線部参照)。これは、単に相手の意見に賛同するだけでなく、自らの意見を主体的に表明し、対話の中で多様な視点を提示し合う姿勢が育まれたことを示している。

#### 例11: 絵画Bの観賞(事後タスク)

1. R: This they mm? mum? the mum?
2. T: Ah mum [surprise]
3. R: Is [gesture: hug] hug
4. T: Ahhh! うんうん
5. R: To the.. woman
6. T: Ohh
7. R: Very sad
8. T: But I think
9. R: Yes
10. T: This woman is ghost
11. R: Ohh! I agree with you

別のやり取りでは次のような場面もあった。

#### 例12: 絵画Bの観賞(事後タスク)

1. H: Mmm.. is this man sleep?
2. N: Ahh sleep! Ohh
3. H: He is closing eyes
4. N: Ahh nice ahh ahhh but in my opinion this [laugh] man is tired
5. H: Ahh

これらの例からは、絵画に対する多様な見方や感想を持つことが自然であるという認識が生徒の間に定着し、それを言語的に表現する手段も獲得されていることが読み取れる。異なる意見が対話の中で受け入れられ、尊重されることで、やり取りはより多面的かつ協働的な意味構築へと発展している。こうした変容は、生徒が単なる感情的同意を超え、より深い理解や多様な視点を言語化し、相互に共有できる認知的・言語的成長を遂げたことの表れであると言える。

#### 4.3.3 自発的な質問の増加

さらに注目すべき点として、事後タスクにおける自発的な質問の増加が挙げられる。事前タスクでは、与えられた質問①～③にとどまる傾向が強かったのに対し、事後タスク(例13)では生徒たちが積極的に追加の質問を投げかける姿が見られた。例えば、What's this?のように、絵画の細部や特徴について具体的に質問を投げかけることで、観察の深まりや解釈の多様化を促進していたことも特徴的である。

#### 例13: 絵画Aについての自発的質問(事後タスク)

1. N: What's what's this?
2. H: Stick
3. N: Ah mm [pointing]
4. H: Ha- hat
5. N: Hat ah hat?
6. H: 帽子
7. N: Ahh sorry

この例では、Namiが相手に絵の細部を尋ねること、単なる情報の受け渡しにとどまらず、相手の反応を受けて確認や修正が行われるなど、相互的な意味構築が展開されていることが分かる。こうした質問のやり取りは、生徒が対話に主体的に関与し、互いの理解を深めていることを示している。

さらに、事後タスクでは Why do you think so? のような理由を掘り下げる追質問も多く見られ、相手の発言の根拠や背景を探ろうとする協働的な対話が成立していた。次のやり取りは、一見すると軽いやり取りに見えるが、その背後ではパートナーの発言に繰り返し理由を問いかけ、思考のプロセスや背景を共有し合おうとする意味構築のプロセスが観察される。例えば以下の例では、Suzu は理由を尋ねる質問を繰り返し投げかけることで、Aya の思考を促していることが分かる。

#### 例14: 絵画Cについての自発的質問(事後タスク)

1. A: But, only one is black
2. S: Why?
3. A: Why why
4. S: Why why do you think so?
5. A: Why do you think so? [laugh]
6. S: Why [laugh] why do you think so?
7. A: Ohh
8. S: One black one black point why do you think so?
9. A: Ummm.. sunny?
10. S: Ahhh
11. A: Or? Err [laugh]
12. S: Earth
13. A: Earth
14. S: Earth バーン バーン [laugh]
15. A: [surprise] [laugh]
16. S: [laugh]
17. A: Earth is break

Suzu の質問(line 2-6)に対して、Aya は一時的に驚く反応を示すものの(line 5)、その後、自己の発言の根拠を考察しながら答えを導き出している(line 9)。さらに Suzu は、言語的には未熟な部分を含みつつも、Aya の考えを補完し発展させる働きを見せ(line 12-14)、最終的に Aya がそれを言語化しまとめる協働的な意味構築の流れが形成されている(line 17)。これらのやり取りは、単なる確認や繰り返しの枠を超え、発話者の思考を深めると同時に、相手との相互補完的なコミュニケーションを促進するものである。こうした探究的かつ協働的な対話の成立には、授業内で行われたスキャフォールディングが大きく寄与していると考えられる。

#### 4.4 発話の文構造や内容表現の変容(仮説4)

VTS 活用授業を通して、生徒の発話における構文的多様性にも顕著な変化が見られた。事前タスクでは、名詞句や単語列、断片的な表現にとどまる発話が多く、文としての構造が成立していない例が目立った(例 2・5・7・9 参照)。例えば、絵画の内容を説明する際に three people, mountain といった単語を並べるのみで、述語を伴わない発話が多く見られた。このような断片的な発話は、視覚的要素の単純な指摘や感覚的印象を伝えるには一定の効果があるものの、発話者の意図や解釈の背景、因果関係などを明確に伝えるには限界がある。また、こうした単語列中心の発話は聞き手の理解を十分に支援せず、対話の深まりや発展を阻む可能性もあった。

一方、事後タスクでは、生徒の発話には主語と述語を備えた完全文が増加し、文同士を接続詞で結ぶ複文も多く見られるようになった(例 3・4・8・10・11 参照)。特に and, because といった基本的な接続詞を活用し、主張と理由を一続きの発話として論理的に構成する力が向上している。以下の例は、その典型的な一例である。

#### 例15: 絵画Bの観賞(事後タスク)

1. K: Ahh I see, I this.. this looks.. like to sad sad depressed
2. E: [nod]
3. K: Because ah the woman is um.. not smiling
4. E: Yeah [nod]
5. K: And.. and um.. animals?
6. E: Yes ahhh [questioning]
7. K: mmmm?
8. E: This is.. I looks like demon?
9. K: De ahhh!

このやり取りにおいて、Kokiは自身の観察や感情を文の形で述べた後(line 1)、接続詞becauseを用いて具体的な理由を説明している(line 3)。さらにandを使って別の観察対象を追加し、対話を発展させている(line 5)。授業で導入された表現パターンを自然に用いることで、視覚的な印象から因果関係の説明へと話の展開をスムーズに行っている点が特に注目される。これは、単なる単語列の羅列から、論理的かつ段階的に意味を構築する文レベルの発話へと成長したことを示している。

こうした構造的な変化は、単語列中心の発話から、論理展開を含む文レベルの発話への移行を示している。主張と理由を結び付ける構文の使用は、聞き手にとって情報の理解を容易にすると同時に、話し手自身が観察や解釈を体系的に整理する助けとなる。このことから、授業で行った段階的な言語的スキファールディングが、生徒の構文的多様性を広げ、論理的な発話を可能にしたと考えられる。

### 4.5 VTS活用授業の成果と考察

本研究で観察された理由付けの自発性、創造的・推測的思考の深化、協働的な対話の成立、さらには文構造の多様化といった一連の変化は、単なる英語学習活動や従来型のペアワークのみでは十分に説明しにくい側面を含んでいる。これらの成果は、VTSを活用した授業実践に帰因し得ると考えられる。

第一に、VTSの特徴である三つの質問枠組みは、生徒に「観察・根拠・解釈」という思考のプロセスを促すものであり、単なる描写や自由発話とは異なり、理由付けや推測を自然に引き出しやすい仕組みを備えている。実際に事前タスクでは質問がなければ根拠提示が困難であった生徒が、授業後には自発的に理由付けを行うようになっていたことは、VTSの枠組みの効果を示唆している。

第二に、題材として用いた芸術作品は、一義的な正解がないため、自由な推測や多様な視点を生みやすい。従来の英語授業で扱う文章読解や教科書教材では、事実確認型のやり取りにとどまることが多いのに対し、VTSでは生徒が自由に推測・創造を行い、互いの意見を補完・対比し合う協働的対話が促進されやすいと考えられる。この点が、創造的・推測的思考の発展に寄与したと考えられる。

第三に、本実践ではVTSの質問枠組みをペア活動として取り入れたことで、生徒同士が互いの発言を引き出し合う形の対話が促された。教師主導の一問一答形式とは異なり、生徒は相手の発言を受けて自ら問い返したり、同意や異論を述べたりする機会が多く与えられた。その結果、発話が単なる観察の羅列にとどまらず、協働的に組み立てられる傾向が見られた。さらに、このようなやり取りを様々な絵画を用いて反復的行ったことから、目標語彙や表現の使用と定着を促進し、協働的対話の成立と構文的発展の双方に一定の寄与があったと考えられる。

以上の点から、本研究で確認された発話の質的变化は、芸術作品を題材とし、VTSの質問枠組みとスキファールディングを統合的に取り入れた授業実践によって促進された可能性が高いと言える。他の方法では必ずしも観察されにくい「根拠提示」「推測の論理化」「協働的対話」「構文的発展」が一貫して確認されたことから、開発したVTS活用授業が生徒の思考力・表現力の発展に有効に寄与する可能性が示唆された。

## 5 本研究の振り返り

本研究の目的は、高等学校の英語授業において、CLIL の枠組みを基盤とし、VTS 活用授業を開発・実践した授業が生徒の思考力と表現力にどのような変容をもたらすかを分析することで、その授業の評価・検証を行うことである。そのために、CLIL の枠組みを基盤とし、英語が苦手な生徒でも取り組みやすいよう段階的な思考的・言語的スキュフォールディングを組み込んだ授業を開発・実践し、英語による絵画鑑賞タスクにおける事前・事後の発話を質的に比較した。

分析の結果、まず思考力の変容としては、理由や根拠を伴う説明が増加し(仮説1)、創造的・推測的な発話が多く見られるようになった(仮説2)。特に絵画の背後にある状況や登場人物の感情を推測するなど、単なる描写を超えた解釈が展開されていた。一方、表現力の変容としては、協働的な意味構築を伴う発話が増え、賛同表現の多様化や異なる意見の提示、理由を求める質問や絵画の細部に関する問いの増加(仮説3)が、発話の継続性と相互理解の促進につながっていた。さらに、これにより、生徒同士が互いの考えを深め合い、主体的に思考を広げる学習環境が形成されていたことが示唆される。また、名詞句や断片的な発話中心だった事前タスクに比べ、事後タスクでは主語＋述語を含む完全文や、接続詞を用いた複文の使用が増加し、主張と理由を一貫して提示する論理的な発話が多く見られた(仮説4)。以上のことから、開発した VTS 活用授業は生徒の思考力と表現力の双方の発展に寄与していることが示された。

しかしながら、本研究にはいくつかの限界も存在する。まず、VTS が対話型であり、生徒同士の共同構成的な活動であることから、個々の生徒の思考力や表現力の変容を厳密に分離して評価することが難しい点が挙げられる。また、特定の高等学校に限られた生徒層を対象に質的分析を中心に行ったため、結果の一般化には慎重さが求められ、発話の頻度や正確さなどの量的指標との連携が不足している。さらに、実際の運用面では、授業モデルの実施には約10時間を要するため、通常授業においてはその時間を確保することが難しいことが予想される。加えて、評価方法についても、より簡便で継続的に活用できる手法の開発が必要である。

今後は、本研究で明らかになった発話の変容を踏まえ、教育現場での持続的な運用方法を検討する必要がある。授業モデルについては、美術高校以外の学校でも運用できるよう、短縮版や年間を通した実施など、より柔軟な形に改良することが望まれる。特に、思考力や表現力の発達を授業中に継続的かつ客観的に把握するためには、評価方法の整備が不可欠である。本研究で用いた分析観点(理由・根拠の明示、創造的・推測的発話、協働の意味構築、構文的多様性)を軸にルーブリックを作成し、評価に活用することが一つの有効な方法となり得る。このルーブリックは、生徒の成長を可視化するだけでなく、生徒自身の自己評価や次の学習目標の明確化にも寄与しうる。また、VTS 活用授業を美術専攻以外の生徒層にも適用し、発話の変容パターンや構文的多様性の発達過程を比較検証することで、その効果や汎用性をさらに検証していくことが望まれる。さらに、発話の質的变化と生徒の自己評価やメタ認知的気づきとの関連を探ることで、VTS を用いた CLIL 型授業の学習効果をより多面的に明らかにすることが期待される。

### 謝辞

本実践研究の貴重な機会を与えてくださいました公益財団法人 日本英語検定協会の皆様、ならびに選考委員の先生方に深く御礼申し上げます。特に、指導助言を賜りました関西大学の竹内理先生には、多大なるご支援と的確なご指導をいただき、心より感謝申し上げます。また、本研究助成のご紹介および構想段階からご助言を頂戴した千葉大学の石井雄隆先生に厚く御礼申し上げます。さらに、絵画鑑賞を用いた CLIL 実践の発想や授業実践に関する貴重なご助言をくださいました福岡教育大学の白井龍馬先生に感謝の意を表します。加えて、VTS 実践に関わる具体的にご指導をいただいた京都芸術大学アート・コミュニケーション研究センターの研究員の皆様、大阪府立港南造形高等学校の先生方にも深く御礼申し上げます。最後に、本授業実践にともに取り組んでくれた生徒の皆さんに心より感謝いたします。

なお、本報告書に関して、開示すべき利益相反関連事項はありません。

## 引用文献

- Aspden, T., Egan, J. P., Bye, L., & Petersen, L. (2022). Using Visual Thinking Strategies to support development of pharmacy student competency in person-centered care. *American Journal of Pharmaceutical Education*, 86(3), 217-225. <https://doi.org/10.5688/ajpe8607>
- Chagall, M. (1912). *La Mariée* [Oil painting]. Centre Pompidou, Paris, France. Retrieved from <https://www.marcchagall.net/la-mariee.jsp>
- Cook, V. (2016). *Second language learning and language teaching* (5th ed.). Routledge.
- Courbet, G. (1854). *Bonjour Monsieur Courbet* [Oil painting]. Musée Fabre, Montpellier, France. Retrieved from Wikimedia Commons.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.
- 池田真. (2011). 「第1章 CLILの基本原理」渡部良典・池田真・和泉伸一『CLIL(内容言語統合学習)上智大学外国語教育の新たな挑戦 第1巻 原理と方法』(pp.1-13) 上智大学出版.
- 池田真. (2016). 「CLIL活用のための新コンセプトと新ツール」池田真・渡部良典・和泉伸一『CLIL(内容言語統合型学習)上智大学外国語教育の新たな挑戦 第3巻 授業と教材』(pp.1-29) 上智大学出版.
- Ikeda, M. (2021). "How would you like your CLIL?": Conceptualising Soft CLIL. In M. Ikeda, S. Izumi, Y. Watanabe, R. Pinner, & M. Davis (Eds.), *Soft CLIL and English Language Teaching* (pp. 9-26). Routledge.
- Kandinsky, W. (1923). *Composition VIII* [Oil painting]. Solomon R. Guggenheim Museum, New York, NY, United States. Retrieved from Wikimedia Commons.
- 文部科学省. (2018). 『高等学校学習指導要領(平成30年告示)』[https://www.mext.go.jp/content/1407073\\_09\\_1\\_2.pdf](https://www.mext.go.jp/content/1407073_09_1_2.pdf)
- 大阪府教育庁. (n.d.). 『大阪版 CAN-DO リスト』[https://www.pref.osaka.lg.jp/documents/9106/leaflet\\_inside1.pdf](https://www.pref.osaka.lg.jp/documents/9106/leaflet_inside1.pdf)
- Poirier, T. I., Newman, K., & Ronald, K. (2020). An exploratory study using Visual Thinking Strategies to improve undergraduate students' observational skills. *American Journal of Pharmaceutical Education*, 84(4), 451-458. <https://doi.org/10.5688/ajpe7600>
- Visual Thinking Strategies. (n.d.). *Founders*. VTS. <https://vtshome.org/founders>
- Yenawine, P. (2013). *Visual thinking strategies: Using art to deepen learning across school disciplines*. Harvard Education Press.

第37回 研究助成

**B** 実践部門・報告Ⅱ・英語能力向上をめざす教育実践

## 教科横断型英語指導による audience awareness と語彙の育成 — 家庭科との横断を通して —

研究者:鳥居 敦子 大阪府／大阪教育大学附属池田中学校 教諭

《研究助言者:和泉 伸一》

### 概要

本研究の目的は、教科横断型英語学習により、(1)生徒のライティングにおける audience awareness が向上し、かつ、別ジャンルのライティングにもその audience awareness は転移するのか、(2)生徒の語彙の幅は向上するのか、を検証することである。大阪教育大学附属池田中学校の中学1年生が6か月にわたり、教科横断型英語学習によるライティングの授業に参加した。生徒の事前・事後の意識調査やライティング課題を分析した結果、教科横断型学習は教科内容を自然に統合でき、現行カリキュラムの枠内でも十分実践可能であることが明らかになった。また、アンケートや作文分析からは、生徒の audience awareness に肯定的な変容が見られ、他者を意識した語彙の使い方や構成の工夫が増加した。さらに、一部の生徒では語彙の多様性にも伸びが認められた。これらの結果は、教科横断型英語指導が生徒の表現力の向上に有効であることを示唆している。

### 1 はじめに

21世紀における言語教育は、単なる知識伝達や技能習得にとどまらず、言語を通じて他者と意味を共有し、関係性を築く能力の育成を目的としている。英語教育においても、学習者が「伝えたい内容を、自らのことばで、相手に届くように」表現する力を育てることが重要視されている。とりわけ、英語を使って「誰に」「何のために」伝えるのかという目的意識を持ち、読み手・聞き手を意識して表現を工夫する姿勢、すなわち audience awareness の育成は、今後の英語教育における中心的課題の一つである。

日本の英語教育も、そうしたコミュニケーション重視の方向へと大きく転換しつつある。文部科学省による中学校学習指導要領(平成29年告示)外国語編では、「言語活動の充実」を図るとともに、「実際のコミュニケーションにおいて、目的や場面、状況に応じて適切に情報や考えを伝える力」の育成が目標とされている。なかでも「書くこと」の領域では、「読み手に取って読みやすく分かりやすい文章が書けるようになる」ために、「受け手を意識し、状況設定を明確にした上で、自分の考えや気持ちを、基本的な語句や表現を用いて簡単な文で書くこと」が求められており、単なる言語知識の運用にとどまらず、読み手を想定して構成や表現を調整する力の育成が重視されている(文部科学省、2017)。

しかし現状では、そうした指導が十分になされているとはいえない。令和5年度全国学力・学習状況調査(中学校英語)の報告では、多くの生徒が自由記述形式の問題において、読み手を意識した構成や表現の工夫に乏しく、情報が羅列的であったり、書き手自身の意図が読み取りにくい文になっている傾向があることが指摘されている(文部科学省、2023a)。これは、文法や語彙の知識がある程度備わっていても、それを「誰かに伝える」という目的で使う力、すなわち audience awareness が十分に育成されていない

ことを表していると考える。

この背景には、教科書におけるライティング活動の課題も影響していると考えられる。例えば、三省堂『New Crown』1年生に掲載されている「ロボットコンテストへの応募」や「English Camp への申し込み」などの活動は、状況設定はあるものの、実際の読み手が詳細に明示されていない。そのため、生徒が「誰に向けて」「どのように伝えるか」を意識しながら言語表現を調整する経験は限定的になりやすい。こうした点においても、教科書中心の活動だけでは audience awareness の育成が十分とは言えない現状がある。

このような課題に対して注目されているのが、「教科横断的英語指導」である。これは、英語科の学びを他教科の内容と結びつけて実施する授業形態であり、英語を実用的かつ意味のある文脈の中で用いることができる点に特徴がある。他教科の教科内容を英語で扱うことで、学習者は単なる英語の練習ではなく、「他者に伝えるための英語使用」を経験することができる。このような実践は内容教育と言語教育を統合しようとする CLIL (Content and Language Integrated Learning) に近いアプローチとして紹介されることもある (Richards & Schmidt, 2013)。本研究では特に CLIL 理論に基づくのではなく、日本の中学校における教育課程内で可能な「教科横断的英語授業」として位置づけ、公立中学校でも現実的かつ実践的な指導法を探ることを目的とする。

本研究は、こうした教科横断的英語指導を中学1年生のライティング活動に取り入れることで、生徒の audience awareness がどのように育成されるかを明らかにすることを目的とする。中学1年生は、特に「書くこと」に関しては英語学習を本格的に開始する初期段階にあり、基礎的な語彙や文構造を習得しながらも、まだ「言語の目的使用」や「読み手を意識した表現」に慣れていない段階である。しかし、「書くことは相手に意味を届ける行為である」という認識を初期段階から育てることで、その後の英語学習に対する姿勢や自己表現力の形成において極めて重要な意味を持つと考える。

本研究では、ある特定の教科(家庭科)との連携を通して、学習者が「誰に向けて何を伝えるか」を明確にしながら英文を書く活動を設計し、指導と実践を行う。そして、その活動の前後における生徒のライティング内容や言語表現の変化、また事前・事後の意識調査(アンケート)の質的ならびに量的データをもとに、生徒の audience awareness の育成プロセスを分析する。これにより、教科横断的な文脈における英語ライティング指導の効果と可能性を実証的に検証し、今後の中学校英語教育における指導法の提案につなげることを目指す。

## 2 先行研究

### 2.1 教科横断型英語教育について

教科横断型英語教育 (Cross-curricular English instruction) とは、英語科と他教科の内容を統合し、異なる教科の知識や技能を英語を通じて扱うことで、言語使用の目的性と実用性を高める教育アプローチである。この方法は、学習者が英語を「学ぶ対象」ではなく「学びの手段」として体験することで、より深い思考や表現につなげることを目指している。富永 (2023) によると、外国語の授業に他教科の題材を取り入れることの有効性はこれまでも語られており、例えば Widdowson (1978) は、現実体験により構成された題材を教えることで、英語を学ぶ意義を伝えやすく、モチベーションも高まるとしている。日本においても、教科横断的な英語指導の可能性は近年注目を集めている。特に、CLIL の手法を使って教科横断的英語指導をする「社会科×英語」「家庭科×英語」などの事例が紹介されている (和泉他, 2024)。

また、学習指導要領においても「教育の内容を教科等横断的な視点で組み立てる」ことが求められており、関係教科と連携して言語活動の質を高める必要性や、特に「言語能力の向上を図る観点から、外国語科など他教科等との関連を積極的に図り、指導の効果を高めるようにすること」が強調されている (文部科学省, 2019)。このため、教科横断的英語指導は、英語を通じた思考力・表現力の育成を支援する手法として、

中学校段階のライティング指導で特に効果を発揮すると考えられる。しかしながら、こうした教科横断型英語指導の実践によって「相手に伝える意識」、すなわち audience awareness がどのように育成されるかに注目した実証研究は、まだ十分に蓄積されていない。

## 2.2 audience awareness について

audience awareness とは、学習者が「誰に向けて」「どのような目的で」書くかを意識しながら、読み手に配慮した内容構成や表現を行う能力である (Hyland, 2005)。Hyland は、ライティングは単なる自己表現ではなく、読み手とのインタラクションの中で意味を構築していく社会的行為であると述べ、書き手が読み手を意識してスタンスやエンゲージメントの工夫を行うことの重要性を強調している。audience awareness は、単なる語彙・文法の正確性にとどまらず、読み手にとって分かりやすく、納得のいく構成や語り口を意識する認知的・社会的スキルであり、上位レベルのライティング能力に直結する (Flower, 1979)。Roen と Willey (1988) も、audience awareness が生徒の文章に与える影響についての研究の中で、文章を書く際に、書き手に対して、読み手がどのような人間であるか、あるいは、読み手に何を伝えるか、読み手が何を知っていて何を知るべきかを考えるように教示した場合は、それらについて考えさせない場合よりも、文章全体の質が高いことを明らかにしている。第二言語としての英語の研究においても、例えば Cumming (1989) は第二言語の書き手の熟達度が上がるにつれて、自己中心的な表現から読み手中心の構成へと変化していくことを示しており、指導において audience awareness を初期段階から意識づけることの意義を強調している。また、日本においても大学生を対象としてではあるが、Sakata (2015) は、読み手意識が英作文の質と量にどのような影響を及ぼすかを検討し、読み手を意識していた学生の方がエッセイの中で具体的な提案を行うことを報告している。例えば、「桜(花見)」の説明では、読み手意識を持つエッセイでは、視覚的イメージや花見の詳細な様子(食べ物や飲み物、家族との交流など)が描かれる一方、読み手意識のないエッセイでは「桜が美しい」といった一般的な記述にとどまる傾向が見られたとしている。そしてその結果、読み手を意識していた学生の方が、より多くの語数を書く傾向があることが明らかになった。

近年では、この audience awareness の育成が、単にテキストの内容や構成にとどまらず、語彙使用の幅 (lexical diversity, vocabulary richness) にも影響を及ぼす可能性が指摘されている。例えば、Liao (2020) は、中国語を学習する英語母語話者を対象に記述作文を分析し、メタ談話的特徴と語彙的・統語的な側面との関連を検討した。その際に、Liao は Hyland (2005) のメタ談話枠組みを援用し、メタ談話を interactive と interactional の二側面に分類している。特に interactional な側面では、自己言及 (self-mention) や 読み手への働きかけ (engagement markers) が含まれ、学習者がどのように読み手を意識して文章を構成しているかを測定する指標とした (Hyland, 2005; Liao, 2020)。結果として、語彙の多様性は、条件・仮定マーカーや読み手への働きかけの表現、そしてメタ談話表現の正確な使用と正の関連があることが示され、統語的複雑さは、フレームマーカーや読み手への働きかけの使用と結びついていたことが報告されている。これらの結果から、Liao (2020) は、読み手を巻き込む表現の活用が語彙の多様性や統語的複雑さの向上と並行して発達する可能性を示唆している。

日本の英語教育においては従来、語彙や文法の正確さ、文構造の形式的習得に重点が置かれがちであり、読み手を意識したライティング指導は十分に行われてこなかった。しかし、学習指導要領 (2017) では、中学校段階でも、「目的や場面、状況を考慮して、自分の考えを簡単な文で伝える」ことがライティングの指導目標として明示され、audience awareness の指導の必要性が示唆されている。

特に中学校1年生の段階では英語の文法的運用能力はまだ初歩的であるものの、伝えたい思いや意図は豊かに持っていることが多い。したがって、この段階で「読む相手がいる」「伝える目的がある」という前提のもとで英語を書く経験を積むことは、英語学習の動機づけやコミュニケーション能力の基盤形成にとっても極めて有効であると考えられる。

このように、audience awarenessの育成は、ライティング能力の発達を支える基盤ではあるが、特に初学者段階の学習者に対してどのような指導が有効であるかについての知見は未だ十分とは言えない。

さらに、Taylor (2021) は、学生が商品レビュー、ブログ、ハウツー記事、旅行ガイドといった異なるジャンルに取り組む際、いずれにおいても「読み手を想定し、その期待やニーズを考慮する」という点が鍵になると繰り返し指摘している。例えば、商品レビューでは「読み手がどのような情報を求めているか」を推定し、ブログでは「対象となる読み手の興味・関心を踏まえて内容や文体を調整する」ことが強調される。ハウツー記事では「読み手の前提知識や技能を考慮した明確で段階的な説明」が不可欠とされ、旅行ガイドにおいても「対象となる読み手(家族連れ、専門職など)に応じた情報と視覚的表現の工夫」が求められる。このように、各ジャンルに固有の特徴を学ぶ過程で、共通する audience awareness のスキルが形成され、他のジャンルにおいても活用され得ることが示唆されている。

## 2.3 研究課題

先行研究からは、教科横断型英語指導が学習者にとって意味のあるコンテキストでの言語使用を可能にし、内容理解や語彙学習において一定の効果があること、またライティングにおける audience awareness は、ジャンルや目的を明確化する指導を通じて育まれ、その効果はジャンルを越えて転移する可能性が示唆されている。しかしながら、中学校の初学年である1年生を対象として、教科横断型の英語指導が生徒のライティングにおける audience awareness や語彙使用にどのような影響を与えるかを検証した実証研究は限られており、特に日本の公立中学校におけるカリキュラム内での具体的実践に関する報告は乏しいのが現状である。

そこで本研究では、中学校1年生を対象に、英語と他教科(本研究では家庭科)を連携させた教科横断型英語指導を実施し、その指導が生徒の英語による表現力のうち、特に audience awareness の育成と語彙使用の変化に与える影響について検証する。

上記目的を達成するために、以下の二つのリサーチクエストション(Research Questions)を設定する：

**RQ1** 教科横断型英語教育は、中学1年生の英語によるライティングにおいて、audience awarenessの向上に寄与し、別ジャンルのライティングにもその audience awareness は転移するのか。

**RQ2** 教科横断型英語教育は、生徒の語彙の使用の幅を促進するのか。

# 3 方法

以下では、研究対象、教材、期間・頻度、手順、データ収集と分析方法について述べる。

## 3.1 研究対象

本研究の対象は、大阪教育大学附属池田中学校の1年生142名(ただし、データ分析は学級閉鎖等の影響を受けずにデータが収集できた80名)である(年齢:12~13歳)。全生徒は、週4回の授業で教科書『New Crown English Series 1』(三省堂)や補助教材を通じて、基本的な文法や語彙、簡単な表現の運用を学んでいる。このうち「書くこと」の授業は平均して2週間に1回の頻度で行われており、本研究の調査は通常の授業の一環として実施された。

## 3.2 教材

英語科の授業では、前記検定教科書を基盤としながら、研究の目的に応じてオリジナルのワークシートやプリントを作成した。特に教科横断型学習を実践する家庭科との連携においては、調理実習に関連した語彙リスト、レシピ構成シート、評価ルーブリック (audience awareness を含む観点) 等を英語で作成し、生徒が「相手に届く」英語使用を意識できるよう設計した。

## 3.3 期間・頻度

本研究の指導およびデータ収集は、2024年9月～2025年2月の6か月間にわたって実施された。英語の授業は2週間に1回の頻度で行われ、限られた授業時間内で教科横断型英語指導は段階的に導入され、計画的なライティング活動が行われた。

なお、勤務校は研究校という位置づけのため、データ収集の同意については生徒の入学時に取得している。

## 3.4 手順

### 3.4.1 プレライティング(2024年9月)

研究開始時に、生徒は「16歳になって海外でホームステイを1年間行うという状況設定の下で、ホームステイ先の家族への自己紹介」をテーマに英作文を書いてもらった。この自己紹介の課題では、教員が生徒にホストファミリーの構成や背景情報を詳細に示し、家族の職業や趣味、性格、年齢など多面的な情報が与えられた。具体的な情報は以下のとおりである。

Goal: 自分のことを知ってもらい、ホームステイ先で自分も相手も心地よく暮らす

Role: 高校1年生(カナダで留学を1年間する予定の人)

Audience: ホストファミリー 家族構成(カナダ在住)

Host father: Tom 45歳 農園で働く、趣味は釣りとスキー

Host mother: Beth 40歳 会計士、趣味はスキーとテニス、料理が得意

Host brother: Dave 17歳 高校生、科学者になりたい、理科が得意、趣味はアイスホッケー

Host sister: Emma 2歳 手がかかる、遊んでほしい、お菓子が好き

Pet: Ben (犬) 大型犬、散歩が必要

Situation: 初めてのホストファミリーへの手紙

そこで生徒がどこまで具体的な相手像を想像しながら文章を作成するかを、書かれた文章の内容や表現からその程度を分析した。また、この段階で Google フォームを用いて audience awareness に関する事前意識調査を実施し、生徒の「読み手を意識した言語使用」に対する意識を測定した。

### 3.4.2 教科横断型英語指導の実践(2024年10月～12月)

英語科と家庭科の連携により、計6回(2週間に1回の頻度)の授業で教科横断型授業を実施した。中心活動は「英語でのレシピ作成と調理実習」であり、具体的な手順は以下のとおりに設計した。

- (1) 語彙・構文の導入: 材料・分量・手順等を表す英語表現の指導
- (2) 実習と言語化: 調理実習の体験を踏まえ、「相手に伝わる」レシピを英語で作成
- (3) 相互評価(peer feedback): 読み手が理解しやすい文章となっているか、語彙選択や説明の丁寧さなど audience awareness の観点から生徒同士でのフィードバックを実施
- (4) 改稿(revision): フィードバックを踏まえてレシピを改稿し、読み手配慮を強化。

この過程で、生徒が具体的な読み手(避難所にいる高齢者、食物アレルギーのある子どもなど)を想定し、その人々の状況やニーズに応じて語彙選択・情報の順序・説明の丁寧さを工夫せざるを得ない状況を意図的に設定した。以下は、具体的な実施内容である。

#### (1) レシピ表現の理解と協働作成(英語授業)

英語教科書『New Crown English Series 1』のレシピ作成に関連する単元(Further Reading Enjoy Sushi)を活用し、生徒に基本的なレシピ表現や構成の特徴を指導した。

続いて、6人グループに分かれて、協働学習として「カレーのレシピ」を英語で作成した。グループ内での話し合いを通じて、各家庭のレシピを参考にしながら、適切な語彙や手順を検討し、文章を完成させた。その際、教員はレシピで使用する命令文の文構造の説明や、Cut, Peel, Pourなどの基本的なレシピで使用する語句の導入を行った。生徒は役割分担として、案を出す人と英語のレシピを書く人に分かれ、グループの中には自分でレシピを書いていない生徒がいるように設定した。こうすることで、後で調理実習をする際に、初見で調理をする人が「読み手」の立場から気づきを得られるようにした。

#### (2) 調理実習①(家庭科授業)

6人グループで作成した自分たちのカレーの英語レシピをもとに、家庭科の授業で実際に調理を行った。6人グループの中で、3人は事前にレシピを実際に協働して書いた生徒、3人は日本語で案のみ出して、英語のレシピライティングにはかかわっていない生徒である。なお、ほとんどの生徒は家庭でカレーの調理を一切したことがなく、レシピを見ながら調理するのが初めての生徒もいた。そこで、事前に家庭で練習することを禁止し、レシピを参照しながら調理を体験し、言語表現の実用性や分かりやすさについて実感を得ることを促した。その結果、例えば、あるグループは単に「にんじんを切る」と書いてだけで何センチ幅に切るかを書いておらず、大切りされた半生の固いニンジンを食べることになったり、「これってどういう意味?」とグループの中で書いた生徒に質問をしている読み手側の生徒もいた。以下は、生徒が振り返りの際に書いた抜粋例である。

- 肉を切る工程を書き忘れたので食べづらくなった。具体的な切り方や大きさ、どのくらい煮込むのかなど、細かい情報をたくさん書いておけばカレーを作ったことのない人にも伝わると思う。
- 図やイラストで説明を補足しておけばさらに良くなったかもしれない。
- 野菜を切るときに、レシピに切り方を書いていなくてよく分からなかった。バタバタしてしまったので、情報が不十分にならないようにしてほしい。

#### (3) レシピの書き直し(英語授業)

調理体験後、グループごとに振り返りの時間を設け、調理を進める中で気づいた課題や改善点について話し合った。改善点としては、「作り方を知らない人が見て作っても分かりやすいように細かいところまで表現をする」「レシピを読む人のことをもっと考えられるようになりたい」「語彙はchopよりもcutの方がよかった」「cutだと分かりにくいから、sliceとかminceとか具体的に書きたい」などが出た。特に、手順の明確さ、語彙の適切性、読み手に伝わる表現の工夫に焦点が当てられ、次のレシピ作成、ならびに調理実習では、読み手を意識して英語でのレシピづくりの際の参考とすることが話し合われた。そうすることで、単なる形式的な言語使用から、読み手を意識した効果的な表現への意識の高まりが期待された。

#### (4) バッククッキング用レシピの作成

修正したカレーのレシピを参考に、次に防災教育の一環として、災害時の「バッククッキング」をテーマにしたレシピ作成に取り組んだ。本校では安全教育に力を入れており、総合的な学習の時間での防災学習を軸にしながら、家庭科と英語の授業で「震災後、水やガスなどのインフラの利用が制限される中で住んでいる地域の全ての住人が利用できる、温かい食事であるバッククッキングのレシピづくりの考案」を

行っている。地域に住んでいる住人は、日本人だけでなく、英語でのレシピがあれば、より多くの人が災害下を食の面で乗り切ることができる。

そこでここでは、6人グループで、グループごとに対象となる読み手を設定した。読み手は、全て英語が理解できる海外の人を対象としつつ、グループによって、糖尿病の持病を持つ人、高血圧等の持病を持つ人、幼児を抱える親、宗教上特定の食材を避ける人(イスラム教)、宗教上特定の食材を避ける人(ユダヤ教)、郷土料理を食べたい人(中国)などを設定した。そして、読み手の条件やニーズに応じたレシピ作成を行った。この過程で、生徒は読み手の健康状態や食の制限、例えば、「食材制限」「味つけの強弱」「食感や安全性」などを考慮し、レシピの内容を考えた。そして、それに沿った形での言語選択が求められた。また、幼児を抱える親が対象であれば、幼児も一緒に調理ができるようにできるだけ絵を入れたり、シンプルな言語を選択するようにしていた。また、アジア圏の人が対象だと、英語を第一言語としない場合があるので、できるだけ簡単な語彙を選択するようにしていた。各グループは料理の内容は6人で考えたが、実際に書く際には3人ずつに分かれて二つのレシピを書いた。これは、後に、自分が書いていない方のレシピを実際に調理してもらうことで、読み手側に立つためである。以下が、生徒が書いたレシピの例である。

#### Porridge

Ingredients: rice 40g, water 230ml, carrot 15g, kiriboshi radish 10g, soy sauce 6g, a dash of salt

1. Put rice into the heatproof plastic bag. Make sure not to close the bag.
2. Put the kiriboshi radish into the bowl with water. Then, throw away the water, after that, refill the bowl.
3. Peel the carrots, then cut them mince.
4. Cut the kiriboshi radish mince.
5. After that, put the vegetables of 3 and 4 into each of the six bags.
6. Put salt and soy-sauce into each bags.
7. Warm it up for thirty minutes.

#### (5)調理実習②(家庭科授業)

各グループが作成したバッククッキング用のレシピに基づき、2回目の調理実習を実施した。6人グループは、3人ずつに分かれており、ここでも生徒が自分たちの英語レシピの実用性を確認するとともに、調理を通じて伝わりやすさや表現の改善点の実感を得るように促した。

#### (6)最終レシピの書き直し(英語授業)

調理実習②の経験を踏まえ、振り返りとして、再度レシピの英語表現を見直し、修正を加えた。

### 3.4.3 ポストライティング と 事後調査(2025年2月)

冬休みを経て3学期には、再び「ホームステイ先への自己紹介」をテーマに英作文を書いてもらった。今回もプレライティングとは異なるホストファミリーの構成や背景情報を詳細に示し、家族の職業や趣味、性格、年齢などの多面的な情報が与えられた。背景情報は以下のとおりである。

Goal: 自分のことを知ってもらい、ホームステイ先で自分も相手も心地よく暮らす

Role: 高校1年生(イギリスで留学を1年間する予定)

Audience: ホストファミリー 家族構成(イギリス在住)

Host father: Richard 45歳 レストランの経営者、趣味は釣りとラグビー

Host mother: Margaret 40歳 薬剤師、趣味はガーデニングとテニス、料理が得意、

Host brother: Alex 17歳 高校生、医者になりたい、得意科目は科学、趣味は音楽鑑賞

Host sister: Brenda 2歳 手がかかる, 遊んでほしい, 甘いお菓子が好き

Pet: Cindy(犬) 大型犬, 散歩が必要

Situation: 初めてのホストファミリーへの手紙

同時に, Google フォームで事後の audience awareness に関する意識調査を実施し, データを量的 (Likert スケール)・質的 (自由記述) の両側面から分析した。

### 3.5 データ収集と分析方法

#### 3.5.1 データ収集

本研究では以下の2種類のデータを収集し, 生徒の audience awareness の変容と, 教科横断型英語教育を事前と事後の間に挟むことによって起こるであろう, ジャンル間での audience awareness の転移の可能性を探る。

##### (1)事前・事後アンケート(Google フォーム):

- 読み手意識に関する自己評価(Likert スケール)
- 自由記述による振り返り

アンケート項目については独自に作成したものであり, 具体例は結果のセクションで示すこととする。

##### (2)ライティング課題

自己紹介文の比較

- プレライティング:ホストファミリーへの手紙①(2024年9月実施)
- ポストライティング:ホストファミリーへの手紙②(2025年2月実施)

同一ジャンル内でのプレ・ポストの比較を通して, audience awareness の成長を質的に検討する。

#### 3.5.2 分析方法

ライティングのデータとアンケート結果について, 以下のように分析を行う。

##### (1)アンケート分析:RQ 1に関する audience awareness の分析

定量データ(Likert スケール)についてはExcelを用いて平均値と標準偏差を算出。自由記述については質的内容分析を行い, 記述の中に現れる「相手意識」「表現の工夫」「困難の克服」などの主題を抽出し, それを事前と事後のライティングで比較した。

##### (2-1)ライティング分析①:RQ 1に関する audience awareness の分析

Hyland(2005b)の Stance & Engagement モデルを参考に, 読み手に向けた表現(例:呼びかけ, 共感, 命令文の丁寧さなど)をカテゴリー化し(表1), 書き手が読み手との関係性をどのように構築しているかを評価し, プレライティングとポストライティングの変化を量的ならびに質的に分析した。

以下が設定した四つの言語的特徴カテゴリーである。

■表1: Stance & Engagement (Hyland 2005b)モデルを参考にした読み手に向けた表現カテゴリー

| カテゴリー                   | 説明            | 例                             |
|-------------------------|---------------|-------------------------------|
| Reader pronouns (読者代名詞) | 読み手を文中に取り込む表現 | you, your, we                 |
| Directives (指示・呼びかけ)    | 読み手に行動を促す表現   | Let's play, Please try        |
| Questions (質問文)         | 読み手に問かける文     | Do you like...?, Is there...? |
| Shared knowledge (共有前提) | 共通の知識に言及する    | As you know, Of course        |

## (2-2) ライティング分析②: RQ 2に関する語彙の幅の分析

英語表現の語彙的な広がりを実量的に評価するため、以下の二つの語彙多様性指標を用いて分析を行う。

- ① Type-Token Ratio (TTR) : 語彙の多様性を簡易に測る指標として、使用された異なった語 (Type) の数を総語数 (Token) で割った。
- ② MTLD (Measure of Textual Lexical Diversity) : 文章の長さに影響されにくい指標として (McCarthy & Jarvis, 2010) , 語彙の繰り返し頻度に基づき、テキストの語彙の広がりを実した。

## 4 結果

### 4.1 事前と事後の Google フォームによる意識調査のデータ

#### 4.1.1 意識調査データ (Google フォーム: 事前・事後)

アンケートでは、以下の設問を提示し、五つの観点について4段階の Likert スケールで自己評価を求めた。

設問: 「次の項目は、まとまった文章を書く時の語彙や表現にどれだけ影響を与えますか」

選択肢: 1: 影響を与えない

2: あまり影響を与えない

3: ある程度影響を与える

4: 大いに影響を与える

調査対象の5項目は、(1)相手の年齢、(2)相手の社会的地位、(3)相手の気分、(4)相手の置かれている状況、(5)自分と相手との関係性、である。これらの項目について、開始時と半年後の平均値を比較した結果を表2に示す。

■表2: ライティング活動における「相手意識」5項目の平均値の変化

| 項目              | 開始時の平均 | 半年後の平均 | 変化量   | 開始時 SD | 半年後 SD |
|-----------------|--------|--------|-------|--------|--------|
| (1) 相手の年齢       | 2.82   | 3.36   | +0.54 | 0.80   | 0.70   |
| (2) 相手の社会的地位    | 2.99   | 3.22   | +0.23 | 0.83   | 0.89   |
| (3) 相手の気分       | 2.77   | 2.76   | -0.01 | 0.90   | 0.88   |
| (4) 相手の置かれている状況 | 3.24   | 3.43   | +0.19 | 0.75   | 0.79   |
| (5) 自分と相手との関係性  | 3.49   | 3.60   | +0.11 | 0.70   | 0.64   |

表2から分かるように、5項目中4項目において平均値が上昇しており、とくに「相手の年齢」と「社会的地位」において顕著な伸びが確認された。さらに、正規性の検定結果に基づき、対応のある  $t$  検定および Wilcoxon 符号付順位検定を選択して実施した。

■表3: 標準偏差と統計的検定結果 (対応のある  $t$  検定 / Wilcoxon 検定)

| 項目              | $t$ 検定 $p$ 値 | Wilcoxon $p$ 値 | 有意差              |
|-----------------|--------------|----------------|------------------|
| (1) 相手の年齢       | 0.0000       | 0.0000         | 有意 ( $p < .01$ ) |
| (2) 相手の社会的地位    | 0.0166       | 0.0096         | 有意 ( $p < .05$ ) |
| (3) 相手の気分       | 0.8688       | 0.9603         | 変化なし             |
| (4) 相手の置かれている状況 | 0.0811       | 0.0807         | 傾向 ( $p < .10$ ) |
| (5) 自分と相手との関係性  | 0.2599       | 0.2000         | 変化なし             |

この結果により、学習者は活動を通じて「相手の年齢」や「社会的地位」といった外的属性を文章を書く上でより重要視するようになったことが窺われる。これは、活動内で読み手の設定を意識させる場面が繰り返し設けられたことと関連していると考えられる。

一方、「相手の気分」については平均値の変化がほとんどなく、統計的にも有意差は認められなかった。この結果は、感情や心理状態といった抽象的要素が学習者にとっては捉えにくく、語彙・表現の選択にまで十分に反映されていなかった可能性や、「気分」はその時々で変化するものでライティングに反映されづらいものである可能性を示唆している。

また、「自分と相手との関係性」も平均値の上昇は見られたが、有意差には至らなかった。

#### 4.1.2 質的分析と考察:自由記述にみる audience awareness の発達と教科横断型学習の影響

生徒が自己紹介文作成時に「意識したこと」について記述した自由回答を対象に、質的分析(Thematic Coding)を実施した。分析の観点としては、本実践の目的である audience awareness の育成に関連する気づきを抽出することを主眼とし、回答に出現する語句・文脈をもとに、七つのカテゴリーを設定した(表4)。なお、自由回答欄の質問は以下のとおりである。

質問:あなたが3学期の自己紹介を書くときに、2学期に行った自己紹介の時よりも意識したことは何ですか。200字程度で書いてください。

その結果、以下のような分類ならびに分析結果となった(回答数はn=80)。

■表4: 事後作文後の生徒の意識調査における、作文作成時「意識した」と回答したものの分類

| カテゴリー               | 出現頻度 | 具体的な記述例(抜粋)                                                                   |
|---------------------|------|-------------------------------------------------------------------------------|
| ①相手の背景理解(文化・知識)     | 34件  | 「海外の人々は日本の文化やマナーを知らないことがあるので、相手に分かりやすいように説明を加えた」「外国の人が読んでも分かるように日本の料理を詳しく書いた」 |
| ②語彙の調整(やさしい表現)      | 32件  | 「分かりやすい単語を使うようにした」「子どもでも分かるような簡単な表現を心がけた」                                     |
| ③相手の状況への配慮          | 30件  | 「ホストファミリーに小さなお子さんがいたので、その子にも伝わるように表現を変えた」「年齢に応じて書き分けた」                        |
| ④内容の選定(相手の興味)       | 8件   | 「相手が好きそうなものを選んで紹介した」「相手との共通点がある趣味を強調した」                                       |
| ⑤形式や見せ方の工夫          | 3件   | 「文字を大きくしたり、区切りを工夫した」「読みやすいように構成を整えた」                                          |
| ⑥メッセージの明確化(一貫性・具体性) | 少数   | 「話がころころ変わらないように工夫した」「具体的に書いてリアリティーを持たせた」                                      |
| ⑦教科横断型英語学習で学んだこと    | 22件  | 「レシピづくりで伝わりやすく書くことの大切さを知った」「家庭科の授業で学んだことを英語にも活かした」                            |

表4から、多くの生徒が自己紹介文作成において、読み手を具体的に想定しながら表現を工夫しようとしていたことが分かる。特に、「相手の背景理解」(34件)や「語彙の調整」(32件)、「相手の状況への配慮」(30件)といったカテゴリーは、audience awareness の中核的要素であり、生徒が言語表現において他者の視点を取り入れる力を獲得しつつあることを示している。

また、読み手の年齢や環境に応じた記述の調整、相手が知らない可能性のある情報を補う説明など、単なる「伝える」から「伝わる」への意識の深化が多く見られた。

### 4.1.3 教科横断型英語学習の影響

前記自由記述(回答数は80件)のうち、「教科横断」「レシピ」「家庭科」「調理」などのキーワードを含む記述を抽出したところ、48件(全体の60%)がこれらの語に言及していた(表5)。

■表5: 事後作文後の生徒の意識調査における教科横断に関する記述の件数と割合

| 項目                | 値     |
|-------------------|-------|
| 自由記述全体数(除外済・空欄除く) | 80件   |
| 該当記述数(キーワードを含む)   | 48件   |
| 該当割合(%)           | 60.0% |

家庭科で行ったパッキングのレシピ作成が英語での自己紹介に影響を与えたことを示す記述も多く見られた。以下は典型的な記述例である。

- 「レシピづくりで、読む人の状況を考えることを学んだので、英語でも相手に伝わるように説明を加えた」

このように、生徒は教科横断を通して、言語表現における他者視点の重要性を実感しており、教科横断で学んだ経験が英語表現に好影響を及ぼしていることが分かる。

### 4.1.4 教科横断学習経験の「深まりレベル」に基づく自由記述の質的階層分析

表4で抽出した、教科横断学習に関する記述のみを対象に、「学びの深まり」を示すレベルを4段階に分類した基準を表6に、その結果を表7に示す。

■表6: 事後作文後のアンケートにおける「学びの深まり」を示す分類基準

| レベル              | 特徴                                   | キーワード例                 |
|------------------|--------------------------------------|------------------------|
| レベル3(抽象的転移・本質理解) | 教科を越えた「伝えることの意義」や、他者視点の本質的理解を言語化している | 「意義」「教科をこえて」「英語にも活かした」 |
| レベル2(具体的応用)      | レシピでの工夫を英語の自己紹介文にも応用している             | 「活かした」「つながった」「同じように」   |
| レベル1(経験の想起)      | IDU での経験や活動内容を単に思い出している              | 「学んだ」「使った」「印象に残った」     |
| レベル0(関連不明)       | IDU に言及しているが、英語表現との関連が不明確である         | —                      |

(注) 本校では「教科横断」の単元をIDU (Interdisciplinary Unit)と呼んでいる。

■表7: 生徒記述の分類結果(定量)

| レベル              | 件数 |
|------------------|----|
| レベル3(抽象的転移・本質理解) | 1  |
| レベル2(具体的応用)      | 3  |
| レベル1(経験の想起)      | 23 |
| レベル0(関連不明)       | 21 |

表7から、全体の半数ほど(47.9%)がレベル1に分類されており、多くの生徒が家庭科と英語での教科横断の学習活動を想起しながら自己紹介文に取り組んでいた様子が窺われる。一方で、英語ライティングにおいて実際に応用している(レベル2)と明確に述べたものは3件、さらに教科を越えて「伝えることの意義」にまで言及した抽象的転移(レベル3)は1件のみであった。これは、学習活動の経験が多くの生徒に記憶されていた一方で、その本質的意味や英語との接続について言語化できた生徒は極めて少数であったことが分かる。

さらに、表8に示す教科横断型学習に関する記述例から、特にレベル2、3に分類された生徒の自由記述において、家庭科での学び(避難所での配慮、読み手の環境、伝わる工夫)が英語表現に応用されていたことが確認された。このような記述は、「他者の視点を持つ」「分かりやすく伝える」などの本質的な表現工夫を、教科を越えて転移させた結果だと思われる。

■表8: 生徒記述の分類結果(代表例)

| レベル  | 記述例                                                                        |
|------|----------------------------------------------------------------------------|
| レベル3 | 「IDUのレシピは避難所という状況下で、どのように見てもらうのかを配慮した。英語でも、読み手の状況を考えて構成を工夫した。」             |
| レベル2 | 「1回目の自己紹介のときは、対象者を一人に絞り、あまり詳しく書いていなかったの<br>で、コミュニケーションの工夫をレシピと同じようにした。」    |
| レベル2 | 「IDUでバッククッキングをしたときに、英語の文でも読みやすさに気をつけて書いた。」                                 |
| レベル1 | 「3学期の自己紹介で意識したことは、文法を正しく使えているかということです。なぜなら、IDUのレシピづくりの時に文法がめちゃくちゃだったからです。」 |
| レベル1 | 「2学期の自己紹介では文法や単語が少なかったけど、IDUを経験してから少しずつ慣れてきた。」                             |

注)本校では「教科横断」の単元をIDU (Interdisciplinary Unit)と呼んでいる。

## 4.2 事前と事後のライティング・データ

ライティング・データの分析結果は、量的、質的に2段階で行われた。以下では、量的な分析結果を最初に述べ、その後で質的な分析結果について報告する。

### 4.2.1 audience awarenessのHylandのEngagement理論による量的分析

中学生がホストファミリー宛に書いた事前と事後の英作文を対象に、Hyland(2005)のEngagementの理論に基づいて、学習者のaudience awarenessの変容を表1で示した表現カテゴリーに基づき、量的に分析した。特に、「Reader pronouns」「Directives」「Questions」「Shared knowledge」の四つの言語的側面に着目し、自己紹介の英作文における読み手との関係構築の程度と性質を比較検討した。

80名分の作文を分析した結果は、表9のとおりである。

■表9: Engagement Markers の出現頻度—事前(Pre)・事後(Post)の比較

| カテゴリー                   | Pre 出現数 | Post 出現数 | 変化量  |
|-------------------------|---------|----------|------|
| Reader pronouns (読者代名詞) | 271     | 527      | +256 |
| Directives (指示・呼びかけ)    | 144     | 220      | +76  |
| Questions (質問文)         | 37      | 191      | +154 |
| Shared knowledge (共有前提) | 0       | 1        | +1   |

分析の結果、Reader pronouns の使用が最も顕著に増加しており (Pre: 271 → Post: 527)、生徒が読み手を直接的に意識する傾向が強まったことが確認された。次いで Questions も大幅に増加しており (37 → 191)、読み手への問いかけを通じた対話的な文体が形成されつつあることが分かる。また Directives についても一定の増加 (144 → 220) が見られ、読み手に行動を促す発話が拡大した。一方で、Shared knowledge の使用はほとんど見られず、Post でわずかに1例が確認されたにとどまった。

#### 4.2.2 audience awarenessのHylandのEngagement理論による質的分析

ここでは、4.1.1のアンケート結果を基に、さらに個々の学習者の変化をより具体的に把握するために、4.1.1で示された意識調査の結果における開始時と終了時のスコアの差に基づき、学習者の audience awareness の伸びを、Positive / Neutral / Negative の三つのカテゴリーに分類した。分類基準は以下のとおりである。

Positive: 終了時の値が開始時より0.2以上高い場合 (意識の向上)

Neutral: 変化量が -0.2以上+0.2未満の場合 (変化なし)

Negative: 終了時の値が開始時より0.2未満低い場合 (意識の低下)

この基準によって、各項目から Positive / Neutral / Negative の代表的な学習者をそれぞれ2名ずつ抽出した。以下に各群に共通する傾向を示し、それぞれの audience awareness の発達段階について考察する。

##### 4.2.2.1 Positiveのカテゴリー

###### (1) Positive Sample 1 (資料1参照)

まずはアンケート結果に基づく audience awarenessの伸びがPositiveであった Sample1によるホストファミリー宛の英作文 (PreおよびPost) を比較分析した結果を表10に示す。

■表10: Sample 1について

| カテゴリー            | Preの使用傾向                                   | Postの使用傾向                                                                                           | 主な変化・特徴                                 |
|------------------|--------------------------------------------|-----------------------------------------------------------------------------------------------------|-----------------------------------------|
| Reader pronouns  | Emma, Beth など個人名への言及はあるが、"you" の使用は文末の1回のみ | Richard, Margaret, Brenda などの呼びかけが増加し、相手の行動や興味に応じた言及が多数                                             | 読み手 (ホストファミリー) の個性を意識した発話が増加している。       |
| Directives       | "Please let me introduce myself." のみ       | "Please get rid of meat from my dish." "I want to cook..." "I would like to teach..." など具体的な提案表現が複数 | 読み手に対する行動要求や希望の表出が増加し、言語の目的意識が明確になっている。 |
| Questions        | 使用なし                                       | 使用なし                                                                                                | 明示的な問いかけはない。                            |
| Shared knowledge | farming, cooking, science などの自己情報を中心       | "Kara-oki is a Japanese pop culture." "Takoyaki is a local food in Osaka." など、文化的情報の紹介が複数出現         | 日本文化に関する知識を読み手と共有しようとする意図が表出している。       |

事前作文では、自己紹介的要素が中心であり、読み手に対して直接語りかける表現や行動を促す表現は限定的であった。一方、事後作文では、読み手の関心や背景を想定した語彙選択が見られ、特定の人物(Margaret, Brendaなど)に向けた行動提案や協働の希望を述べることで、対話的な関係構築の意識が高まっている様子が窺われる。

さらに、kara-oki, takoyakiなどの日本文化に関する情報を文中で共有しようとする姿勢が確認され、Shared knowledgeの活用という点が見られた。これはHyland(2005)が述べる「読み手をテキスト内に巻き込み、共通理解の前提を作る」言語戦略の実践に相当する。

一方で、疑問文などを用いた読み手への能動的な関与(Questions)は未使用であった。

## (2) Positive Sample 2 (資料2参照)

次に、Sample 2によるホストファミリー宛の事前と事後の英作文を比較分析した結果を、表11に示す。

■表11: Sample 2について

| カテゴリー            | Preの使用傾向                                                             | Postの使用傾向                                                                       | 主な変化・特徴                                         |
|------------------|----------------------------------------------------------------------|---------------------------------------------------------------------------------|-------------------------------------------------|
| Reader pronouns  | "you"や"Let's play ice hockey"の使用が見られるが、呼びかけの頻度は限定的                   | Brenda, Cindy, Alex, Margaretなど、相手を個別に指名して語りかける表現が複数登場                          | 読み手を能動的にテキストへ取り込もうとしている。                        |
| Directives       | "Please let me introduce myself." "Let's play ice hockey."などの簡単な呼びかけ | "Please teach me", "Could you take me", "Look at the picture"など、具体的に多様な指示的表現が増加 | 指示表現のバリエーションと文脈適合性が高まり、読み手とのやりとりが活性化している。       |
| Questions        | "Could you take me there?"など依頼型の疑問が限定的に見られる                          | "What anime do you like?" "Do you have any allergies?"など、相手の考えや情報を引き出す明示的質問が導入  | Engagementとしての双方向的問いかけが強化され、interaction性が増している。 |
| Shared knowledge | Niagara Falls, lobster dishなどを挙げてはいるが、読み手と共有する意図は弱い                  | Jagariko, KitKat, Big Ben, British dishesなど、文化的・地理的知識を共有する表現が複数登場               | 日本の文化・食・観光を積極的に紹介している。                          |

Sample 2の事後作文では、事前作文と比較して、読み手(ホストファミリー)への関心や働きかけが質的に大きく変化している。特に、読み手個々人への具体的な呼びかけ(Reader pronounsの使用)、行動を促す指示表現(Directives)、読み手に問いかける疑問文(Questions)の導入といった側面において、テキスト内でのインタラクションの意識が表れていることが読み取れる。

また、日本の文化や食品を視覚的に紹介する記述("Look at the picture"による菓子の説明など)は読み手の関心を引き、共通の知識(Shared knowledge)を介して関係を築こうとする表現行為と評価できる。これらは、Hyland(2005)が指摘する「読み手をテキストの中に巻き込み、共に意味を構築する」engagementの具体的な実践であり、audience awarenessの深化が明確に確認された。

## 4.2.2.2 Neutralのカテゴリーの結果について

### (1) Neutral Sample 3 (資料3参照)

次に、アンケート結果でaudience awarenessの発達がNeutralであったSample 3についてホストファミリー宛の事前と事後の英作文を比較分析した結果を表12に示す。

■表12: Sample 3について

| カテゴリー            | Preの使用傾向                                                         | Postの使用傾向                                                                                     | 主な変化・特徴                          |
|------------------|------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|----------------------------------|
| Reader pronouns  | you, your dog を使用。だが読み手の特性に踏み込む表現は少ない                            | you は継続して使用されているが、呼びかけ対象は一般的な host family であり、特定の人物像を想定していない                                  | 読み手の存在を認識はしているが、内容は一方的である。       |
| Directives       | "Let's go fishing", "Let's play the game", など、自分の趣味に読み手を巻き込む呼びかけ | "Let's play Mario!" のみ。自己の趣味の延長であり、読み手の関心への配慮は見られない                                           | 使用されているが、自己中心的な呼びかけにとどまっている。     |
| Questions        | "What game do you like?" と1文のみ使用                                 | "Do you have a cafeteria at school in the UK?" "Is baseball a famous sport in the UK?" と2文に増加 | 質問数は増加しているが、読み手の文脈を深く想定した問いではない。 |
| Shared knowledge | "I have goldfish", "I like baseball" など自己情報中心                    | "Mario is a popular game character", "Osaka has a famous castle" など文化的共有を意識した情報が追加            | 自己→文化へ範囲は広がったが、文化の紹介にとどまっている。    |

Sample 3の事後作文では、事前作文と比較して、読み手への問いかけ(Questions)や文化的知識(Shared knowledge)が含まれている点で一定の変化は確認できた。しかし、それらの多くは書き手自身の興味関心に基づく発話であり、読み手の属性・関心・期待を想定して表現を調整するという観点は乏しいのではないと思われる。例えば、"Let's play Mario!" や "Do you have a cafeteria?" といった表現は読み手に語りかけているものの、相手の興味や背景を想定したものではなく、結果としてテキストが一方的な情報発信にとどまっている。

また、Shared knowledgeにおいても「大阪城」や「マリオ」など広く知られる情報を挙げてはいるが、それが読み手との相互理解を深めるための戦略としているとは言い難い。このように、Sample 3の事後作文における audience awareness の変化は、表現の表層的拡張にとどまり、言語使用や対話性の観点からは限定的であると思われる。これがアンケート結果における Neutral と一致する点である。

## (2) Neutral Sample 4(資料4参照)

続いて、Sample 4のホストファミリー宛の事前と事後の英作文を比較分析した結果を表13に示す。

■表13: Sample 4について

| カテゴリー            | Preの使用傾向                                        | Postの使用傾向                                                                                   | 主な変化・特徴                       |
|------------------|-------------------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------|
| Reader pronouns  | "Beth, can you make seafood pizza?" という直接呼びかけあり | "Margaret, can you make pizza?" "Alex, let's read it together" など、複数の個別呼びかけ                 | Post では、ホストファミリーの個人名を多く用いている。 |
| Directives       | "Please call me Miki." 程度に限定                    | "Please tell me good sightseeing spots." "Please teach science to me." など、より明確な働きかけが増加      | 指示・依頼の表現がバリエーション豊かに使用されている。   |
| Questions        | "Beth, can you make seafood pizza?" のみ          | "Do you have food allergies?" "Do you have a lot of snow in the UK?" など、読み手の情報を引き出す質問が複数    | 疑問文が明示的に増加している。               |
| Shared knowledge | "I can take care of Emma" など、自己情報中心             | "I heard you are good at science", "We don't have much snow in Osaka" など、読み手と共通の情報に基づく言及が追加 | 読み手に関する事前情報や地理的知識を活用している。     |

Sample 4の事後作文では、事前作文と比較して読み手に対する語りかけの質に明らかに変化が見られる。事前作文では主に自己紹介的な構成にとどまり、語りかけも限定的であったが、事後作文では、読み手の個別性(Alex, Margaret, Brendaなど)を意識した呼びかけ、質問、行動の提案が複数見られた。特に、“I heard you are good at science.”や“Let’s read it together.”など、相手の情報を踏まえて文脈を構成する工夫(Shared knowledge)が見られ、Hyland(2005)が提示するEngagementの実践的応用として評価できる。一方で、表現のバリエーションは増加しているものの、内容の深まりや読み手の反応を想定した表現(例:推測や条件づけ、文化比較など)は依然として限定的である。

以上のことから、Neutralに分類されたSample 3およびSample 4の事後作文では、形式的なaudience awareness発展は確認されたものの、読み手との意味ある関係構築という点では、呼びかけや質問の使用は増加するも、自己中心的または表面的な表現の傾向が見られた。

このように、Neutralグループの作文では語彙や文法の表面的な発展が見られる一方で、テキスト内で読み手との「対話」を構築しようとする意識が限定的である点が、アンケート結果の「変化があったが明確ではない」という評価とも一致した。

#### 4.2.2.3 Negative のカテゴリーの結果について

##### (1)Negative Sample 5(資料5参照)

最後に、audience awarenessの伸びがNegativeであったSample 5によるホストファミリー宛の事前と事後の英作文を比較分析した結果を表14に示す。

■表14: Sample 5について

| カテゴリー            | Preの使用傾向                                            | Postの使用傾向                                    | 主な変化・特徴                                 |
|------------------|-----------------------------------------------------|----------------------------------------------|-----------------------------------------|
| Reader pronouns  | “I would like to walk with Ben” など限定的               | “I want to talk with you all” 程度にとどまる        | 相手への直接の呼びかけは少なく、読み手の個性や関心を意識した表現はなし。    |
| Directives       | 使用なし                                                | “Please visit” のみ (文脈としては観光案内的)              | 呼びかけとしての機能は薄い。                          |
| Questions        | 使用なし                                                | “How about you?” “Do you like Miso soup?” のみ | 一応の問いかけはあるが、必ずしも読み手の関心を想定したものではない。      |
| Shared knowledge | “I’m good at drawing”, “I have a hamster” などの個人情報のみ | “Takarazuka has many good points” などの地域情報追加  | 自己中心的情報提示が中心で、読み手との相互理解を意図した構成にはなっていない。 |

Sample 5の事後作文では、事前作文と比較して、生活習慣(朝食や起床時刻)や出身地(Takarazuka市)についての言及が加わっている。しかしながら、Hyland(2005)の“engagement”の観点から見ると、読み手を「対話者」として取り込もうとする意識は希薄なようである。特に以下の点で、audience awarenessの発展が限定的であることが感じられた。まず、Reader pronounsやDirectivesの使用が最小限であり、読み手に対して直接的に語りかける姿勢が弱いという傾向が見られた。また、Questionsは使用されているが、内容が個人的で突発的であり、例えば、“Do you like Miso soup?”のように、読み手の属性や関心に基づいて問いかけているとは言い難いものである。Shared knowledgeの扱いも一方的であり、自身の出身地や日常を「紹介」しているにとどまり、読み手との関係性を構築する意図は見られなかった。

これらの特徴から、Sample 5の作文は、表現の表層的多様化は進んでいるものの、根本的な「読み手の存在を前提にした書き方」には至っていないと評価される。これは、アンケート結果においても、自己中心的に書いてしまったという回答や、読み手への配慮が十分でなかったという振り返りとも一致していた。

## (2) Negative Sample 6(資料6参照)

次に同じく、audience awarenessの伸びがNegativeであったSample 6によるホストファミリー宛の事前と事後の英作文を比較分析した結果を表15に示す。

■表15: Sample 6について

| カテゴリー            | Preの使用傾向                                        | Postの使用傾向                                                   | 主な変化・特徴                                         |
|------------------|-------------------------------------------------|-------------------------------------------------------------|-------------------------------------------------|
| Reader pronouns  | “your dog”, “your favor” のみ。読み手をテキストに取り込む使用は限定的 | “Let's play tennis”, “if it's ok”, “please tell me” などの表現あり | Post で若干増加しているが、依然として読み手像は曖昧であり、対話的な意図は希薄である。   |
| Directives       | “Please let me introduce myself.” のみ            | “Let's play tennis”, “Please tell me” などが使用される              | 呼びかけは若干見られるが、定型・汎用的なフレーズにとどまる。                  |
| Questions        | “What kind of music do you like?” のみ            | 同じ質問がそのまま繰り返されている                                           | 読み手への問いかけは形式的で、内容や文脈の深まりが見られない。                 |
| Shared knowledge | “I like Namie Amuro” のみ。個人的嗜好の提示                | “I recommend ONE OK ROCK” などが追加                             | 音楽という話題を通じた情報提示はあるが、読み手の趣味や文化背景との接点づくりにはなっていない。 |

表15によると、Sample 6の事後作文では、形式的には“Let's～”や“Please～”といったEngagement表現の使用も見られるが、それらの多くは定型的・表面的な言語使用にとどまっている。特に、“What kind of music do you like?”の繰り返し使用に見られるように、問いかけはあるが読み手の属性や興味を想定した構文とは言えず、応答や対話を引き出す工夫がなかった。また、読み手への呼びかけや依頼も、“Please tell me.”や“Let's play tennis.”といった自己の興味関心に基づいた一方的な提案であり、読み手の関心や反応を想定した書き方には至っていないという傾向が見られた。Shared knowledgeの面でも、音楽やペットに関する話題は出ているものの、読み手と共有可能な話題として戦略的に選ばれた形跡は見られなかった。

このように、Sample 6の英作文は自己紹介的な情報提示の枠を出ておらず、読み手との関係を構築する言語的試みは限定的である。これはアンケートで示された「読み手意識の変化が乏しい」という評価と一致している。

## 4.3 事前・事後の語彙データ

### 4.3.1 英作文における語彙の幅の量的分析

学習者の英作文能力の発達を定量的に把握するために、語彙の幅に着目し、二つの代表的な語彙多様性指標であるType-Token Ratio(TTR)とMeasure of Textual Lexical Diversity(MTLD)を使用した。まずは、Text Inspectorを用いて、全体(80人分)の事前(Pre)・事後(Post)の自己紹介文を対象に語彙の変化を測定した。結果を表16に示す。

■表16: TTR・MTLD要約統計( $n = 80$ )

| 指標        | 平均      | 標準偏差   | 中央値     | 最小値     | 最大値     |
|-----------|---------|--------|---------|---------|---------|
| TTR_Pre   | 0.6403  | 0.0704 | 0.6419  | 0.4354  | 1.0000  |
| TTR_Post  | 0.5746  | 0.0714 | 0.5707  | 0.4194  | 1.0000  |
| TTR_Diff  | -0.0657 | 0.0605 | -0.0617 | -0.2144 | +0.0917 |
| MTLD_Pre  | 86.93   | 32.80  | 85.50   | 0.00    | 151.00  |
| MTLD_Post | 88.03   | 33.50  | 81.38   | 0.00    | 180.00  |
| MTLD_Diff | +1.10   | 41.21  | -5.00   | -87.00  | +116.25 |

まず、TTRの結果として、事前作文の平均値は0.6403( $SD = 0.0704$ )、事後作文の平均値は0.5746( $SD = 0.0714$ )であり、平均値は約0.066ポイント減少した。中央値も同様に低下しており、全体的に語彙の多様性が減少する傾向が確認された。最小値から最大値の範囲を見ても、事前作文では0.435から1.000であったのに対し、事後作文では0.419から1.000であり、個人差はあるものの、全体として減少傾向にあることが明らかである。

次に、MTLDの結果では、事前作文の平均値は86.93( $SD = 32.80$ )、事後作文の平均値は88.03( $SD = 33.50$ )であり、平均値ではわずかに上昇した。しかし中央値は85.50から81.38へとむしろ低下しており、学習者間のばらつきが大きいことが窺える。MTLDの変化量(Diff)の平均は+1.10であったが、標準偏差は41.21と大きく、-87.00から+116.25まで幅広い変化が見られた。

これらの結果から、TTRでは語彙の幅が全体的に減少しているのに対し、MTLDでは平均値がほぼ横ばいで、個人差が大きいという特徴が明らかとなった。

#### 4.3.2 抽出した Sample 1～Sample 6の英作文における語彙の幅の分析

続いて、アンケートにおける事後評価としての audience awareness との関係を見るため、4.2.2 であげた、6名の生徒(Sample1～6)の事前と事後の自己紹介文を対象に語彙の変化を測定した。その結果を表17に示す。

■表17: 抽出したSample1～6の事件事後自己紹介文の語彙の幅分析

| Sample   | TTR_Pre | TTR_Post | MTLD_Pre | MTLD_Post | MTLD変化量 |
|----------|---------|----------|----------|-----------|---------|
| Sample 1 | 0.59    | 0.55     | 45.63    | 48.32     | +2.69   |
| Sample 2 | 0.67    | 0.52     | 69.77    | 55.41     | -14.36  |
| Sample 3 | 0.59    | 0.57     | 42.86    | 56.41     | +13.55  |
| Sample 4 | 0.64    | 0.61     | 69.70    | 59.22     | -10.48  |
| Sample 5 | 0.66    | 0.63     | 56.40    | 85.19     | +28.79  |
| Sample 6 | 0.72    | 0.60     | 90.89    | 78.57     | -12.32  |

6名の生徒のうち、Sample 1、Sample 3、Sample 5の3名はMTLDが上昇し、英語表現において語彙の多様性が向上したことが認められる。一方で、Sample 2、Sample 4、Sample 6の3名はMTLDが減少した。Sample 3やSample 5の作文では、新たなトピック(例:地域紹介、文化比較、献立や活動の提案など)が増え、それに応じて語彙の幅が大きく拡張された。これにより、MTLD値が顕著に向上した。一方で、Sample 4やSample 6の作文では、使用される語彙が似通っていたり、表現の繰り返しが多かったりする

傾向が見られた。このことがMTLDの数値にマイナスの影響を及ぼしたと考えられる。以下に、詳しく分析を述べる。

### (1) 語彙の多様性が向上した学習者

Sample 1 (MTLD +2.69, TTR -0.04) は語彙の多様性が若干向上した。TTRの低下は事後作文で文量が増加したことに伴う自然な変化であり、むしろ語彙の反復を効果的に活用していると解釈できる。

Sample 3 (MTLD +13.55, TTR -0.02) はPostで話題の範囲が広がり、新たな名詞、動詞が導入された。TTRの維持とMTLDの大幅な増加から、より多様な語彙を繰り返しなく活用できるようになったと評価される。

Sample 5 (MTLD +28.79, TTR -0.03) は最大の語彙的成長を示した。内容的には日本文化や自然描写に関する抽象的表現が増えており、それがMTLDの大幅な上昇につながっている。

### (2) 語彙の多様性が減少した学習者

Sample 2 (MTLD -14.36, TTR -0.15) は、事後作文で読み手への呼びかけや共感的な語りが増え、語彙の使用が特定の範囲に集中したと考えられる。読み手であるホストファミリーへの呼びかけ(質問、依頼、共感など)が増加し、より読み手に配慮した構成となっている。文の目的や構成が変化したことによって、語彙の繰り返しや限定的な語彙使用が増え、結果としてMTLDが低下した可能性がある。

Sample 4 (MTLD -10.48, TTR -0.03) は、話題の広がりに対して使用語彙の繰り返しが目立ち、特に感情語や質問文型で同一語の反復が増えていた。

Sample 6 (MTLD -12.32, TTR -0.12) は、事前作文で非常に高いMTLD値を示していたが、事後作文では語彙がより実用的な表現に限定されたことで、多様性の観点ではやや減少となった。

### (3) 事前・事後の代表文例から見る結果

では、具体的にどのような変化があったのか。表18に、Sample 1～6の事前・事後の自己紹介の代表例の比較を示す。

■表18: 個別事例における作文比較 (Sample1～6, 事前・事後の代表文例)

| Sample   | Preの使用傾向                                                                                                                                                            | 主な変化・特徴                                                                                                                                                                                                                                |
|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sample 1 | "My favorite character is Miffy." "I'm not good at cooking. So, I want to learn cooking for Beth."                                                                  | "I like Miffy. She is very cute character in the world." "I'm not good at cooking. But, I know Japanese food well. So, I want to cook Japanese food at Richard's restaurant with Margaret."                                            |
| Sample 2 | "I would like to go to Niagara Falls." "Could you take me there?"                                                                                                   | "I would like to go to Big Ben." "Could you take me there on weekend?"                                                                                                                                                                 |
| Sample 3 | "My hobby is fishing." "Tom, let's go fishing with me."                                                                                                             | "My hobby is playing games." "Let's play Mario! Mario is a popular game character."                                                                                                                                                    |
| Sample 4 | "I like seafood pizza and grapes. Beth, can you make seafood pizza for me?" "I like watching ice hockey. I would be grateful if I could play ice hockey with Dave." | "I like pizza very much. My favorite pizza is margherita pizza. Margaret, can you make pizza? I'm expecting it." "My favorite Japanese anime is Case Closed. I can tell about the interesting story ... Alex, let's read it together." |
| Sample 5 | "I'm good at draw pictures." "I would like to draw pictures with Emma."                                                                                             | "I'm good at drawing the landscape in the UK." "I would like to cook British food with you."                                                                                                                                           |
| Sample 6 | "My favorite singer is Namie Amuro." "I especially like her song 'Tempest'."                                                                                        | "I recommend ONE OK ROCK." "They are rock band musician."                                                                                                                                                                              |

表18から見る記述内容の変化については以下のことが言える。

#### Sample 1

MTLD(+8.92), TTR(+0.07)と改善が見られた。事前では“Miffy is very cute.” “I like Miffy.”といった短い嗜好表現にとどまっていたが、事後では“Miffy is popular in Japan.” “She has many goods.”といった情報追加や文化的参照が含まれるようになった。語彙選択も“cute”から“popular” “goods”へと広がり、単なる個人的嗜好から情報共有的な語りへ発展している。

#### Sample 2

数値的にはMTLD(-14.36), TTR(-0.15)と低下が見られたが、事後作文では“I would like to go to Big Ben on weekend.”といった時間条件や依頼表現が導入され、読み手(ホストファミリー)への呼びかけや共感的な語りが増加していた。これは、語彙使用が特定の範囲に集中し、繰り返しが生じた結果、指標上は多様性が減少したと言える。

#### Sample 3

MTLD(+6.41), TTR(+0.04)と数値が向上した。事前は“I like soccer.” “Soccer is fun.”といった繰り返しが多かったが、事後では“I like soccer because I play it with my friends.” “My favorite team is Gamba Osaka.”など、理由付けや固有名詞の導入によって表現が精緻化した。語彙の幅は“fun”や“play”といった汎用語から“team” “friends”など具体的な社会的語彙へと展開しており、読み手に自分の経験を伝えようとする姿勢が窺える。

#### Sample 4

MTLD(-10.48), TTR(-0.03)と数値は微減した。事前では“seafood pizza” “ice hockey”といった多様な話題を展開していたが、事後では“margherita pizza” “Case Closed”などに話題を絞り、その魅力を繰り返し強調する傾向が見られた。特に感情語や質問形式の表現が増えたため、同一語彙の反復が多くなり、数値的には語彙多様性が減少した。

#### Sample 5

MTLD(+5.18), TTR(+0.05)と増加を示した。事前は“I like reading.” “My hobby is reading.”といった抽象的な嗜好表現が中心であったが、事後には“I like reading Case Closed.” “I read it every week.”など具体的作品名や習慣表現が用いられるようになった。

#### Sample 6

MTLD(-12.32), TTR(-0.12)と比較的大きな減少が見られた。事前は“Namie Amuro”など嗜好を列挙する形で多様な語彙が使われていたが、事後では“I recommend ONE OK ROCK.”といった推奨表現が中心となり、表現が機能的かつ実用的に限定された。そのため数値的には多様性が下がったが、読み手への働きかけというコミュニケーション機能の発展が見られる。

このような結果から、英語学習初期段階においても、生徒によって語彙の多様性の変化パターンが大きく異なることが明らかとなった。特にaudience awarenessを強く持った作文では、語彙の多様性よりも語用的機能(相手に配慮した繰り返し、明確な語彙選択)が優先される傾向が見られた。

## 5 考察

### 5.1 研究のねらいの再確認

本研究では、中学1年生を対象に、英語と家庭科を連携させた教科横断型のライティング指導を行い、英語による自己表現において audience awareness がどのように育成されるかを検討した。また、それに付随して語彙の使用にどのような変化が見られるかについても分析を行った。

### 5.2 読み手意識の変化とその背景(RQ 1に関連して)

本研究のアンケート結果と作文分析から、生徒は活動を通じて「相手の年齢」や「社会的地位」といった外的属性を文章に反映させる意識を高めていたことが分かった。読み手を意識した書き方を模索するようになってきたことが窺われる。自由記述では、「文化的背景への配慮」「語彙の調整」「相手の状況に応じた工夫」といった観点が多く挙げられていた。

さらに、作文分析においても、Reader pronouns, Directives, Questions の使用が大幅に増加しており、生徒が読み手を文中に巻き込み、対話的な文体を形成しようとする傾向が表れた。特に事後作文において、“Do you have any allergies?” “Margaret, please tell me...”といった表現が見られたことは、単なる情報提示にとどまらず、読み手を意識した相互作用的構成へと発展しつつあることが示唆される。特に、audience awareness がポジティブに変化した生徒の作文では、相手に語りかける表現や文化的情報を共有しようとする姿勢が見られ、書き手が「読まれること」を意識して表現を調整している様子が見てとれた。一方で、読み手を意識しているように見えても、それが表面的なものにとどまっている場合もあり、相手との関係性や感情の理解といった、より深いレベルでの audience awareness は、まだ十分には育っていないことも示された。これは、表現幅やスタイルの問題だけでなく、学習者自身が他者の視点に立つて物事を考える経験がまだ限られていることとも関係していると思われる。

また、本研究は ジャンル間の転移についても一定の成果があったと考えられる。すなわち、家庭科の「レシピ作成」を通じて学んだ「相手に伝わる表現の工夫」が、英語の自己紹介作文に応用されていた点である。アンケートの自由記述の分析からは、家庭科での学びが英語の表現に影響を与えている様子が見られた。「レシピづくりのときに、読む人の状況を考えたことが英語にも活きた」といった記述は、学習者が経験を他の教科に応用していることを示している。この結果は、Taylor (2021) が指摘する「読み手意識のスキルはジャンルを超えて転移可能である」という主張を支持している。ただし、本研究において深いレベルで転移できていた生徒はごく少数であり、大多数は「経験の想起」にとどまった。したがって、今後はより一層の転移を促すために、本研究で行ったような学習経験を積ませることが必要である。その上で、教員による、他ジャンルとの比較検討のような明示的なメタ認知的指導も必要になってくるのではないかと考えられる。

### 5.3 RQ 2:教科横断型英語教育は、生徒の語彙使用の幅を促進するのか。

本研究で用いた語彙多様性指標(TTR・MTLD)の分析からは、全ての生徒に一律な改善が確認されたわけではなかった。むしろ、一部の生徒においては、audience awareness の向上に伴い、特定語彙の反復使用が増え、数値的には語彙多様性が低下する現象も見られた。これは、読み手との関係性を強化しようとする際に、“you” “please” “let’s” などの限定的な表現が繰り返されやすくなるためであり、語彙の幅よりも読者適合性を優先した結果と解釈できる。一方で、一部の生徒の作文には、新たに導入された語彙の使用が確認された。具体的には、事後作文において“allergies” が読み手への質問表現に登場し、ま

た自由記述では「文化や食べ物を詳しく説明した」「相手が知らない可能性のある日本の習慣を紹介した」といった言及が見られた。これらは、レシピ作成や防災教育を通じて獲得した「相手の状況に応じた語彙・内容の工夫」が自己紹介作文に反映された事例である。つまり、語彙そのものの量的な拡張よりも、語彙の文脈的適合性や相手に応じた選択が進んでいたと評価できる。

これらの結果は、学習者が教科横断的な文脈を通じて獲得した語彙を、別ジャンルの作文にも活用できる可能性を示している。Sakata(2015)が示した「読み手を意識した学生の方が具体的な語彙を増やし、より多くの語数を書いた」という知見とも部分的に一致している。なお、質的变化には個人差があり、英語学習初期段階においても、生徒によって語彙の多様性の変化パターンが大きく異なることが明らかとなった。特に audience awareness を強く持った作文では、語彙の多様性よりも語用的機能(相手に配慮した繰り返し、明確な語彙選択)が優先される傾向が見られた。これらは、相手に応じて適切な語彙を選び、それを意味のある形で構成する力を身につけたと窺われる。

以上を総合すると、語彙の「幅」の量的拡張は全体として限定的であったが、読み手に適合する語彙を選択する「質的变化」は明確に見られた。したがって、今後は単純な語彙多様性指標に依拠するのではなく、「相手に応じた語彙使用の適切さ」や「文脈的関連性」を評価できる指標を導入する必要があると考えられる。

## 5.4 教科横断型英語指導の可能性と課題

今回の実践では、英語科と家庭科を連携させることで、「相手に伝える」必要のあるライティング活動を試みた。結果として、英語でのレシピ作成と調理実習を組み合わせることで、生徒が実際の読み手を具体的に想定しながら、語彙や構成を工夫するようになる様子が観察された。こうした教科横断型の活動は、CLIL の考え方にも通じるものであり、教科をまたいで意味のある文脈の中で言語を使う経験が、生徒の学びを豊かにできると考えられる。また、授業準備を進める中で、複数の教科が連携して授業を設計するためには、教科ごとの目標や評価の捉え方の違いを十分に考慮した調整や教材面での工夫が求められることも確認できた。このような苦労は決して少なくはないが、本研究のように英語が単なる「教科」ではなく他教科の学びを深めるための「手段」として活用される経験は、生徒にとって英語学習への動機づけを高める絶好の機会になり得る。実際、調理実習を終えた後に英語表現を見直す生徒の姿からは、言葉が現実の行動とつながっているという実感が育まれていたように思われる。また、調理実習を通してレシピを校正するというプロセスを経る中で、「相手に伝わるためにはどうすればよいのか」を考え、それが後のホストファミリーへの手紙の際にも応用できたと考える。

## 5.5 総合的考察と教育的示唆

本研究から得られた知見は次の3点に整理される。

- 教科横断型英語指導は、中学1年生において audience awareness を高める上で役立つ。
- その効果は、同一ジャンル内にとどまらず、他のライティング課題にも部分的でもある程度転移する。
- 語彙の面では、その幅の数値的な拡張は限定的であるが、読み手を意識した語彙選択という質的側面で改善が期待できる。

このような結果は、教科横断型英語指導が学習者に「言語を社会的に用いる」経験を初期段階から提供し、英語教育におけるコミュニケーション能力の育成に資することを示唆している。

以上を踏まえると、教科横断的英語指導は、日本の中学校における現実的なカリキュラムの中で実践可能であり、生徒の audience awareness の育成にとって極めて有効なアプローチであると結論づけられる。

## 5.6 今後の課題

本研究は、中学1年生を対象に、家庭科との連携を通じた教科横断型英語指導の効果を検証したものである。しかしながら、いくつかの限界が存在する。第一に、本研究の調査対象は特定の中学校(附属校)の1学年に限定されており、地域性や学校文化による影響を排除できない。そのため、本研究の結果を日本全国の中学校に一般化するには慎重さが求められる。

第二に、調査期間が6か月間と比較的短期であった点である。audience awareness の向上やジャンル間転移が長期的に持続するのか、あるいは一過的な効果にとどまるのかは、本研究の範囲では明らかにできない。特に、audience awarenessの発達には相応の経験と時間を要するものと考えられ、生徒のライティング能力は学年の進行とともに発達するため、長期にわたる縦断的な追跡研究が不可欠である。

第三に、語彙の幅に関する分析指標としてTTRとMTLDを用いたが、これらは数値的な多様性を捉えるにとどまり、「読み手への適切さ」「文脈との関連性」といった質的側面を十分に測定できていない。結果として、語彙使用の変化を過小評価している可能性もある。今後は、語彙多様性指標とあわせて、語彙選択の適切性を質的に評価するルーブリックの開発や、自動言語解析ツールとの併用を検討する必要があるだろう。

第四に、生徒の自由記述に基づく質的分析において、教科横断活動を「経験の想起」としてのみ言及するケースが多数を占め、本質的な転移に至った生徒は少数にとどまった。6か月という限られた学習経験であったため、今後は経験の拡張が必要と考えられる。そして、その経験によって徐々に学習が深まり、転移を促すことができるものと考えられ、その過程において、例えばリフレクション・ジャーナル やピア・ディスカッションを組み込み、学習者が自らの気づきを抽象化・一般化できる場を設けることが求められる。

以上を踏まえると、今後の研究の展望としては以下の方向性が考えられる。

- 異なる地域・学校における追試を行い、教科横断型英語指導の汎用性と実践可能性を検証する。
- 継続的な縦断的調査を通じて、audience awareness の発達が学年進行や他ジャンルの学習にどのように影響するかを明らかにする。
- 語彙使用の評価指標を精緻化し、「多様性」と「適切性」の双方を捉える包括的な枠組みを構築する。
- 教科横断的活動のリフレクション支援を強化し、生徒が学びの本質を抽象化して他ジャンルへ転移させるプロセスを支援する。

このような課題を克服することにより、教科横断型英語指導の価値がよりはっきりと見えてくるのではないかと考える。

## 謝辞

この研究の機会をくださった公益財団法人 日本英語検定協会、選考委員の先生方、特に貴重なご助言と的確なご指導をくださいました和泉伸一先生に心より感謝申し上げます。また、研究に参加してくれた生徒の皆様に深くお礼申し上げます。なお、本報告書に関して、開示すべき利益相反関連事項はありません。

## 引用文献

- Cumming, A. (1989). Writing expertise and second-language proficiency. *Language Learning*, 39(1), 81-141. <https://doi.org/10.1111/j.1467-1770.1989.tb00592.x>
- Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. *College English*, 41(1), 19-37. <https://doi.org/10.2307/376357>
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.
- Hyland, K. (2005b) 'Stance and engagement: a model of interaction in academic discourse'. *Discourse Studies*, 7(2), pp173-192.
- 和泉伸一(編著)(2024)『CLILで広がる英語授業』大修館書店
- 国立教育政策研究所(2023a). 「令和5年度全国学力・学習状況調査報告書 中学校英語」[https://www.nier.go.jp/23chousakekkahoukoku/report/data/23meng\\_k.pdf](https://www.nier.go.jp/23chousakekkahoukoku/report/data/23meng_k.pdf)
- Liao, J. (2020). Metadiscourse, cohesion, and engagement in L2 written discourse. *Languages*, 5(2), 25. <https://doi.org/10.3390/languages5020025>
- McCarthy, P. M., & Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42, 381-392.
- 文部科学省. (2017). 中学校学習指導要領:外国語編. [https://www.mext.go.jp/content/20210531-mxt\\_kyoiku01-100002608\\_010.pdf](https://www.mext.go.jp/content/20210531-mxt_kyoiku01-100002608_010.pdf)
- Richards, J., & Schmidt, R. (2013). *The Longman dictionary of language teaching and applied linguistics (4th Ed.)*. Abingdon, Routledge.
- Roen, D. H. & Willey, R. J. (1988). The effects of audience awareness on drafting and revising. *Research in the Teaching of English*, 22(1), 75-88.
- Sakata, R. (2015). The effects of audience awareness on Japanese students' English writing [Master's thesis, University of Yamanashi]. Yamanashi University Repository. <https://yamanashi.repo.nii.ac.jp/record/1274>
- 三省堂. (2021). New Crown English Series 1. 三省堂.
- Taylor, E. (2021). Using online genres to promote students' audience awareness. *Journal on Empowering Teaching Excellence*, 5(2), Article 4. <https://doi.org/10.26077/ac79-ce31>
- 富永裕子(2023)『中学校「英語」×「教科」の試み：教科横断型学習と自己評価』三恵社
- Widdowson, H (1978) . *Teaching language as communication*. Oxford University Press.

## 資料1

### 事前作文

September 10 2024

350 Albert street Suite 600,  
Ottawa, Ontario, K1R 1B1, Canada

Dear Host family,

My name is [redacted]. I'm very happy to stay at your house for a year. Please let me introduce myself.

I'm sixteen years old, and I live in Osaka, Japan. I'm good at swimming. I can sing well. I have a lot of stuffed toys. My favorite stuffed toy is a rabbit, and my favorite character is Miffy. Talking about singers, I'm a big fan of Yoasobi.

My mother works for a part-time job at a school, and my father works for the local government. My maternal grandparents live in Tochigi, and my paternal grandparents live in Hyogo. And, my cousins live in Tochigi too.

I like dogs. But, I'm a little scared. So, I want to get used to dogs. I like little children. I want to play with Emma. I'm interested in farming. I want to help Mr. Tom. I'm not good at cooking. So, I want to learn cooking for Beth. And, my dream is to become a scientist. But, I'm not good at science. So, I want to learn science with Dave in Canada.

I'm sure I will improve my English during my stay. And, I want to play with Ben. I'm looking forward to seeing you soon. Best wishes, [redacted]

### 事後作文

February 17  
8 Springfield Road,  
Bury St Edmunds,  
Suffolk,  
IP33 3AN,  
England

Dear Host family,

My name is [redacted]. I'm very happy to stay at your house for a year. Please let me introduce myself.

I'm sixteen years old. I like to sing songs. So, I like to go to kara-oki. Kara-oki is a Japanese pop culture that people go to a shop to sing. I like pizza and, I don't like meat. So, please get rid of meat from my dish. But, I like hamburger. I live in Osaka, Japan. Takoyaki which is a local food in Osaka is very delicious. My maternal grandparents and my cousins live in Tochigi. Tochigi's strawberry is sweet. These are so many delicious local foods in Japan, and they are all delicious.

I'm not good at cooking. But, I know Japanese food well. So, I want to cook Japanese food at Richard's restaurant with Margaret. My dream is to be a pharmacist. So, I want to learn science from Alex. And, I want to sing the Japanese animation song together because I like it, and perhaps Alex likes too. I like Japanese language. So I would like to teach Japanese to Brenda. I hope she will also like Japanese.

I mean, I have a lot of expectations to do in the UK. I want to play with you. I'm looking forward to seeing you soon. Best wishes,  
[redacted]

## 資料2

### 事前作文

September 3, 2024  
350 Albert street, Suite 600, Ottawa, Ontario, K1R1B1, CANADA

Dear Mr. Tom

My name is [redacted]. Please call me Rika. I'm very happy to stay at your house for one year. Please let me introduce myself.

I'm 16 years old. I live in Osaka with my family. I have one sister who is an elementary school student. Actually, I have a dog, so I can take care of Ben. I'm a member of the tennis club and math club at my school. As I'm interested in science, please teach me about science and English.

I don't have any allergic, and I can eat almost all food. I often spend time by doing babysitting: a baby of my parents' friend, so I can take care of Emma. My hobby is playing ice hockey too. Let's play ice hockey with me, Dave!

I have several things to ask for you. First, I would like to go to Niagara Falls. So, could you take me there? Also, could you make lobster dish for me? I can make Japanese food, so I will make it for you!

I'm sure I will improve my English during my stay. I'm looking forward to seeing you soon.  
Best regards,  
[redacted]

### 事後作文

February 4, 2025  
8 Springfield Road, Bury St Edmunds, Suffolk, IP33 3AN, England

Dear Richard, Margaret, Alex, Brenda, Cindy

My name is [redacted]. Please call me Rika. I'm very happy to stay at your house for one year. Please let me introduce myself.

I'm 16 years old. I live in Osaka, Japan with my family. I have one sister who is an elementary school student. She is very cute. I like a little children like her. So I'm glad that I can stay with 2 year old Brenda. I often spend time by doing babysitting: a baby of my parents' friend, so I would like to take care of Brenda. Actually, I have a dog, so I'm also happy to see your dog, Cindy. I can take care of Cindy if it's OK. I'm a member of the tennis club at school now and I like playing it. So, Margaret, let's play tennis with me!

Also I'm interested in science and I was a member of the science club at my elementary school. Please teach me about science more, Alex. I like rock music and Japanese anime, so I will introduce my favorite music and anime to you. What anime do you like? I want to know about your favorite!

I don't have any allergies, and I can eat almost all food. Do you have any allergies? I will bring Japanese snacks to you. Look at the picture. The left side is [redacted]

Jaganiko. It's like chips. The right side is called KitKat. It's a chocolate bar. They're very yummy!

I have several things to ask for you. First, I would like to go to Big Ben. So, could you take me there on weekend? And, my hobby is eating various foods around the world. Recently I'm interested in British food. So, Richard, could you make British dishes for me? I can make Japanese food. I will make it for you!

See you soon!  
[redacted]

I'm sure I will improve my English skills during my stay. I'm looking forward to seeing you soon.  
Best regards,  
[redacted]

資料3

事前作文

September4,2024

350Albert Street  
Suite600,Ottawa,Ontario,K1R1B1,CANADA

Dear host family.  
My name is [redacted] Please call me Kurokou.  
I am very happy to stay at your house for two weeks from this summer. Please let me introduce myself.  
I am sixteen years old. I live in Osaka Japan with my family. I have goldfish. They are very beautiful.  
My birthday is August 16th.  
My hobby is fishing so Tom, Let's go fishing with me. I like dogs so I want to walk your dog.  
I like baseball and I belong to baseball club. I can play the game very well. Let's play the game! What game do you like?  
I like meat very much. Can I cook meat for you?  
I am not good at speaking English. But I want to be native. I am looking forward to seeing you soon.Best regards.

Thank you!

[redacted]

事後作文

Janually22,2025  
springfield Road,Bury St Edmunes, Suffolk,IP33 3AN

Dear Host Family,

My name is [redacted] Please call me Koki.  
I am very happy to stay at your house for one year.  
Please let me introduce myself.  
I am sixteen years old.My birthday is August 16th 2011.  
I am from Hyogo in Japan but I live in Osaka now.  
Osaka has a famous castle.



My family members are father,mother and sister.  
My father is working for a company.My mother is a nurse.  
My sister is an elementary school student.  
And I have a pet.It is a dog.The dog's name is Charo.  
In the past I had gold fishes.But they were dead.



My school is OKU high school.  
The school has a cafeteria. It has many meals.  
They are very delicious.  
Do you have a cafeteria at school in the UK?



By the way,my hobby is playing games.  
Let's play Mario!Mario is a popular game character.

I belong to a baseball club.  
I practice baseball everyday but it is very difficult.  
Is baseball a famous sport in the U.K?



I am not good at speaking English. But I want to be native. I am looking forward to seeing you soon.Best regards.

Thank you!

資料4

事前作文

September, 17,2024  
352 Albert street,Suite,600 Ottawa,Ontario  
K1R, 1B1,CANADA

Dear Host family

My name is [redacted] Please call me Miki.  
I'm very happy to stay at your house for one year.  
Please let me introduce myself.

I'm sixteen years old.I live in Osaka,Japan with my family.  
I have two younger brothers and I always take care of them.So I can take responsibility for taking care of Emma.  
I like watching ice hockey. I would be grateful if I could play ice hockey with Dave.  
I like seafood pizza and grapes.  
Beth,Can you make seafood pizza for me? I'm expecting it.

I'm sure I will improve my English  
I'm looking forward to seeing you soon.  
Best regards.

[redacted]

事後作文

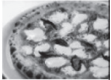
5,Marth,2025

8 Springfield Road,  
Bury St Edmunes,  
Suffolk,  
IP33 3AN

Dear Host family !

My name is [redacted] Please call me Miki !  
I'm very happy to stay at your house for one year in the UK.

I'm sixteen years old.I live in Osaka,Japan with my family.  
We went to some foreign countries before,but I haven't been to the UK.  
So I'm very happy to go to the UK.  
Please tell me good sightseeing spots in the UK.  
I don't have a lot of snow in Osaka. Do you have a lot of snow in the UK?



I like pizza very much.My favorite pizza is margherita pizza.  
Margaret,can you make pizza? I'm expecting it.  
I can make Japanese food for you.

I'd like to bring some Japanese souvenirs.Do you have any food allergies?  
Please,let me know.

Alex,I want to be a doctor,some as you But,I don't like science.  
I heard you are good at science.How wonderful ! If it's possible,please teach science to me!  
And,my favorite Japanese anime is Case Closed.  
I can tell about the interesting story of Case Closed.  
Alex,let's read it together,I'll bring comic books !

I have dogs.She is very difficult to handle.  
So I can walk with your dog,Cindy if you allow me to do.

I have two brothers.They are so noisy.  
So I can take responsibility for taking care of Brenda.  
Brenda?Let's play a game with me !

I'm looking forward to seeing you soon.  
Best regards [redacted]



## 資料5

### 事前作文

September 18,2024

350 Albert Street,Suite 600,Ottawa,Ontario,K1R 1B1,CANADA.

Dear Host family,

My name is [REDACTED]. I'm very happy to stay at your house for one year,please let me introduce myself.I'm sixteen years old.I live in Hyogo,Japan.I have one sister.She is a school child.I have a hamster.I'm good at draw pictures.I would like to draw pictures with Emma.I would like to walk with Ben.My hobby is collecting stationery. I'm sure I will improve my English during my stay.I'm looking forward to seeing you soon.

Best regards,

[REDACTED]

### 事後作文

February 19th

8 Springfield Road, Bury St Edmunds, Suffolk, IP33 3AN

Dear host family,

I'm [REDACTED]. Thank you for accepting me. Let me introduce myself. I am 16 years old and go to OKU high school. I decided to study abroad in England because I want to study English pronunciation. So I would like to talk a lot with you all. I'm good at drawing the landscape in the UK and cooking British food with you. I have three members in my family and have a hamster. We all live in Takarazuka city. Takarazuka has many good points. For example, the scenery is beautiful and you can feel nature. It's very nice place, please visit. Also, I usually get up at 7 o'clock and eat bread for breakfast. How about you? Do you like Miso soup? My mother taught me how to cook Miso soup, and I hope to cook them for you.

I'm looking forward to meeting you.

Sincerely yours,

[REDACTED]



↑ Miso soup                      ↑ Takarazuka city                      ↑ my hamster

## 資料6

### 事前作文

September 20,2024

350 Albest Street,Suite  
600,Ottawa,Ontario,K1R  
1B1,CANADA

Dear Host Family,

My name is [REDACTED].  
I'm very happy to stay at your house for one year.  
Please let me introduce myself.

I'm sixteen years old.  
I live in Hyogo,Japan.  
I have a dog and two cats. I like animals very much!  
So I would like to walk your dog.  
I can play tennis. I'm in the tennis club.  
I often listen to music. My favorite singer is Namie Amuro.  
I especially like her song "Tempest".  
What kind of music do you like?  
Please let me know your favor.

I'm sure I will improve my English during my stay.  
I'm looking forward to seeing you soon.  
Best regards,

[REDACTED]

### 事後作文

February 17, 2025

8 Springfield Road,  
Bury St Edmunds,  
Suffolk,  
IP33 3AN

Dear Host Family

I hope you are doing great. My name is [REDACTED].  
I'm very happy to stay at your house for one year.  
Please let me introduce myself.

I'm sixteen years old. I live in Amagasaki, Hyogo with my family.  
I like to play tennis. Also I'm a member of the tennis club at my school. Let's play tennis together!  
My hobby is listening to music. I recommend ONE OK ROCK. They are rock band musician.  
What kind of music do you like?  
I have two cats. So I would like to walk your dog if it's ok. If there is anything I can do, please tell me.

I'm sure I will improve my English during my stay.  
I'm looking forward to seeing you soon.  
Best regards.

[REDACTED]



# イントネーション指導を統合した音読活動： 発話のComprehensibilityと語彙保持に及ぼす効果

研究者：根子 雄一郎 東京都／東京大学教育学部附属中等教育学校 教諭

《研究助言者：和泉 伸一》

## 概要

本研究では、日本語を母語とする高校生を対象に、明示的なイントネーション指導を統合した音読活動の効果を検証した。リズム・強勢・イントネーションの指導を行った介入群とリズム・強勢のみの指導を行った対照群を比較し、実験1で発話のComprehensibilityへの効果を、実験2では語彙保持への効果を測定した。実験1の結果、両群とも発話のComprehensibilityが有意に向上した。既習の文章の読み上げ課題では両群に有意差は見られなかったものの、初見文章の読み上げ課題では介入群が対照群を有意に上回った。一方、実験2では、低得点層において対照群が介入群と比べ有意に高い成績を示した。これらの結果から、イントネーションの統合的指導は未知の文脈における発話のComprehensibilityの向上に有効である一方、学習者の習熟度や教材の難易度に応じた指導の設計が重要であるという教育的示唆が得られた。

## 1 はじめに

従来の英語発音指導では、訛りを軽減しネイティブスピーカーのモデルに近づくことを目標とするNateness Principle が重視されてきた。しかし近年では、発音の明瞭さ(Intelligibility)や理解しやすさ(Comprehensibility)の向上を目的とする Intelligibility Principle がより重視されており(Derwing & Munro, 2015), 従来型の音素中心の指導に加え、プロソディを重点的に扱うことが推奨されている(鈴木, 2024)。

プロソディの中でも、イントネーションはComprehensibilityに大きな影響を与えるとされ(Pickering, 2001; Kang, 2010; Gilbert, 2014; Derwing & Munro, 2015; Levis, 2018), コミュニケーションにおいて高い重要性を有している。しかしながら、教室において指導が十分に行われているとは言い難い。その理由として、イントネーションの複雑さ(Ramírez Verdugo, 2006)や、指導に対する教師の自信の不足(Gilbert, 2014)が指摘されている。イントネーションの複雑性の一例として、ピッチの種類やそれらの組み合わせによって生じるパターンの多様性が挙げられる。例えば、O'Connor and Arnold(1973)は、ピッチの変化を7種類に分類し、それらの組み合わせから10種類の音調群を提示している。このようなイントネーションの複雑な体系を第二言語学習者が完全に習得することは極めて困難である(Roach, 2009)。そのため指導の際には教授・学習可能性を考慮した上で、適切な目標設定のもと、指導内容の取捨選択をする必要がある(Pennington & Rogerson-Revell, 2019)。

また、英語の授業では語彙や文法など教授すべき内容が多岐にわたり、発音指導に割ける時間が多くないことも指導不足の一因として指摘されている(Rogerson-Revell, 2011; Gordon & Darcy, 2016)。こ

のことから発音指導を独立して行うのではなく、授業内の他の活動に組み込むことを Grant (2014) は提唱している。しかし、具体的な指導の方法が広く共有されているとは言えない現状がある (大和 & 磯田, 2020)。

統合的なイントネーションの発音指導を行う際の有力な活動の一つとして、音読活動が挙げられる。音読は、中学校学習指導要領 (文部科学省, 2018) の「読むこと」の活動の一つとして記載があるように、これまでも中等教育現場で広く取り入れられてきた。音読では、文章を前もって読むことで区切り位置や文強勢の位置を予測するトレーニングが可能であり、イントネーションを指導する上で有用であると考えられる (Rogerson-Revell, 2011)。

以上を踏まえ、本研究では、まず中等教育において指導すべきイントネーションの事項を選定し、その上でイントネーションの発音指導を音読活動に統合する方法を考案する。そしてその効果を多角的に検証することで、指導に関する教育的示唆を提示することを最終的な目的とする。

## 2 先行研究

### 2.1 イントネーションの機能と産出

Roach (2009, p.3) はイントネーションを “the use of pitch of the voice to convey meaning” と定義し、その機能を以下の四つに分類している。

1. 心的態度機能 (attitudinal function): 話者の感情や態度を示す。
2. アクセント機能 (accentual function): 情報の新旧や対比を示す。
3. 文法機能 (grammatical function): 文の種類や句・節の区切りを示す。
4. ディスコース機能 (discourse function): 情報の新旧や発話の継続・完了を示す。

これらの機能を通じて、話者は聞き手に対して意味や意図を伝達する。そして、これらの機能を実現するためのイントネーションの産出プロセスを、Wells (2006) は以下の三つの段階に分けて説明している。

1. トーナリティ (Tonality): 発話内容を意味的・文法的なまとまりに応じて音調群 (tone unit) に区切る。
2. トーニシティ (Tonicity): それぞれの音調群の中で最も情報的に重要な語に音調核 (nucleus) を置く。特定の強調されるべき情報がない場合 (broad focus) は、原則として音調群の中の最後の内容語に核が置かれる (渡辺, 1994; 牧野, 2005)。
3. トーン (Tone): 音調核にピッチの変化を付与し、話者の態度や意図を示す。

このように、話し手がイントネーションの機能を用いて意味を伝えるためには、上記の産出プロセスを経る必要がある。本研究におけるイントネーションの指導は、このプロセスを基盤として構成する。

### 2.2 イントネーションの指導内容

指導すべきイントネーションの要素について、Wells (2006) の産出プロセスに沿って検討する。ここでは指導の対象として、英語の習熟度が中級程度の高校生 (中学校範囲の学習内容を十分に理解している。CEFR A2-B1 を想定) を想定する。まず、Tonality について、音調群は統語的なまとまりと一致することが多く (Wells, 2006; Rogerson-Revell, 2011)、音調群への分割においては文法的知識を直接適用できるため、教授可能性および学習可能性が高いと判断される。また、情報のまとまりを明確にすることは、Comprehensibility の向上に貢献するため (渡辺, 1994)、学習の重要性も高い。

次に Tonicity について、前提として音調核の配置には文強勢の知識が必要となるため、まず内容語と機能語の区別と英語のリズムについて教授する必要がある。その前提を理解すれば、broad focus の場合

には原則に基づく音調核の配置が可能であるため、こちらも教授・学習可能性が高いと考える。さらに、情報の新旧や対比に基づく音調核の配置には、文章の読解スキルが転用できるため、学習負担も比較的小さいと推測される。重要情報の強調や情報の新旧・対比の明確化は情報伝達において重要であり (Low, 2006; Pennington & Rogerson-Revell, 2019)、こちらも Comprehensibility の向上に大きく寄与するため、その学習の必要性は高いと言える。

一方で、Tone については学習内容を絞り込む必要があると考える。第一に、高下降調や低下降調のようにピッチの高低を厳密に区別するトーンの種類は、聞き分けが難しく、学習者にとって習得が困難である可能性が高い (渡辺, 1994)。第二に、音調核の前の部分 (前核部) のピッチの変化についても、地域差 (渡辺, *ibid*) や年代による変化 (Lindsey, 2019) が存在するため、一般化されたルールを提示することが難しい。第三に、トーンを選択による心的態度の表現は、個人差が大きいことに加え (渡辺, *ibid*)、その複雑さに伴う学習負担に対して得られる効果が意思の伝達という観点からは限定的である (Pennington & Rogerson-Revell, 2019) ことから、指導の優先順位を下げるのが望ましい。

これらの点を踏まえ、本研究ではトーンの種類を基本となる下降調、上昇調、下降上昇調の3種類 (Wells, 2006) に絞り、前核部のピッチ変化には言及しないこととする。また、トーンを選択については心的態度機能とは結びつけず、原則に基づいて判断できるよう教授する。具体的には、平叙文、Wh 疑問文、命令文には下降調を、Yes/No 疑問文には上昇調を付与するよう指導する (Wells, *ibid*)。さらに、文の途中にある音調群の核には、下降上昇調が使われる頻度が高いことから (渡辺, 1994; 牧野, 2005)、これを付与するように指導する。

## 2.3 イントネーションと音読活動の統合的指導

Celce-Murcia, Brinton, and Goodwin (2010) は、コミュニケーション的な発音指導のための体系的枠組みを提示している。この枠組みでは、まず対象とする発音項目について明示的に説明し、学習者の意識を高めるとともに、音声インプットを通して学習対象の音声の特徴に関する気づきを促す (Description and Analysis / Listening Discrimination)。次に、音声の形式面に焦点を当てた反復練習を行い (Controlled Practice)、その後、形式と意味の双方に注意を向けさせる練習へと発展させる (Guided Practice)。最終段階では、意味の伝達を中心としつつも音声面への意識を維持する活動を行う (Communicative Practice)。つまり、この枠組みでは、学習者の意識の高揚から形式面の習得、さらに意味に焦点を当てた活動へと段階的に移行する指導プロセスを提案している。なお、Celce-Murcia et al. (*ibid*) は、これらの段階が必ずしも直線的に進行するものではなく、学習項目や学習者のニーズに応じて段階間を往還することを認めている。

本研究では、この枠組みを基盤として、イントネーションと音読活動を統合した指導法を検討する (表 1 参照)。第1段階では、学習者の意識を高めるために、イントネーションの機能と産出プロセスについて明示的な説明を行い、音声の意味および形式的特徴への意識を高める。第2段階では、教師による文章のモデル音読を提示し、学習者に音調群の区切りや音調核におけるピッチ変化への気づきを促す。第3段階では、教師のモデルに続くコーラスリーディング (鈴木 & 門田, 2012) を通じて、音声の形式面に注目した音読練習を行う。第4段階では、文章の意味に焦点を当て、その意味を効果的に表現するために音読練習を行う。最終段階では、文章内容に関連したロールプレイやリテリング活動を実施し、意味重視のやり取りの中で形式面への意識を維持させる。

■表1: イントネーションと音読活動の統合的指導案(Celce-Murcia et al., 2010を参考に作成)

| 段階                          | 活動                     | 留意点                  |
|-----------------------------|------------------------|----------------------|
| 1. Description and Analysis | イントネーションの明示的説明         | 音声の意味・形式に対する意識の高揚    |
| 2. Listening Discrimination | 教師のモデル音読のリスニング         | 音声の形式に対する気づき         |
| 3. Controlled Practice      | 教師のモデルに続くコーラスリーディング    | 音声の形式の習得             |
| 4. Guided Practice          | 文章の意味を表現するための音読練習      | 文章の意味と音声の形式の結びつけ     |
| 5. Communicative Practice   | 文章の内容に関連したロールプレイやリテリング | 意味重視の中で音声形式に対する意識の維持 |

上記のイントネーション指導を統合した音読活動は、音声面の習熟を促すだけでなく、音読そのものの効果を高める可能性がある。音読には言語材料の記憶と内在化を促す効果があり(竹内, 2018)、実際に日本人高校生を対象に音読の英語表現の習得への効果を検証した高橋(2007)では、様々な種類の音読を繰り返し行った群が、数回のみ音読を行った群と比べ、英語表現の定着度が高かったことが報告されている。効果的な音読のためには、文章の意味内容と文法構造を十分に理解した上で行うことが重要であり、その過程で文章中の言語材料が内在化され、長期記憶への転移が生じるとされている(門田, 2012)。

この点において、文を音調群に分割する作業は、意味的・文法的まとまりへの意識を促す効果が期待される。また、音調核の配置を検討することにより、情報の新旧や対比といった情報構造への意識が高まり、文章の意味内容の理解がより深まる可能性がある。Craik and Tulving(1975)によれば、深い水準での意味処理が行われることが語彙の長期記憶保持につながるとされるため、このような統合的音読活動は語彙の記憶保持にも有効である可能性があり、その効果の検証には意義があると考えられる。

## 2.4 イントネーションの指導の効果

イントネーションの習得が明示的な指導なしには困難であることは、先行研究によって示唆されている。例えば、斎藤(2018)による日本語母語話者の大学生を対象とした調査では、英語圏での10か月間の留学経験を経た後も、音調核の配置に関して顕著な進歩が見られなかったことが報告されている。また、筆者が中学生を対象に行った調査では、明示的なイントネーション指導を実施せず、教科書に掲載された文章のモデル音声を繰り返し聞かせた上で、音読練習を実施したところ、モデル音声と同様の音調核の配置を再現できた学習者は全体の4割に満たなかった(Nego, 2024)。

一方で、明示的な指導がイントネーションの習得を効果的に促進することも、多くの研究から明らかにされている。Derwing, Munro, and Wiebe(1998)は、カナダ在住の成人学習者を対象に、セグメンタルおよびプロソディに焦点を当てた指導の効果を検証した。この研究では、プロソディの指導を受けた群が、文の読み上げ課題と即興の発話課題において、Comprehensibilityを有意に向上させたと報告されている。Zhang and Yuan(2020)は大学生の中国人英語学習者を対象に、Derwing et al.(ibid)と同様のデザインのもと実験を行った。結果から、プロソディの指導を受けた群が文の読み上げ課題と即興の発話課題において、Comprehensibilityを有意に向上させたことが明らかとなった。また、Ramirez Verdugo(2006)は、スペイン語を母語とする大学生を対象に、イントネーションの明示的な説明と可視化されたピッチ曲線をフィードバックとして与える指導を行った。その結果、学習者の発話のIntelligibilityが有意に向上していることが示された。また Saito and Saito(2017)による研究では、大学生の初級レベ

ルの日本人英語学習者に対してプロソディの指導を実施した。その結果、文章の読み上げ課題において Comprehensibility が有意に向上し、さらに yes/no 疑問文と wh 疑問文の発話において適切なピッチの変化を付与することができるようになった。

上記の研究が示すように、プロソディの指導は概して発話の Comprehensibility を向上させることを示唆している。しかし、Derwing et al. (1998) と Zhang and Yuan (2020) では、具体的にどのようなイントネーションの指導が行われたかは言及がない。Ramirez Verdugo (2006) は、イントネーションの明示的な説明を行ったことには言及があるが、具体的にどのイントネーションの要素について説明が与えられたのか述べられていない。Saito and Saito (2017) ではイントネーションの指導の内容を説明しているが、文の種類に応じたピッチの選択のみで、音調群への分割や音調核の配置については指導していない。また、Derwing et al. (ibid), Zhang and Yuan (ibid), Saito and Saito (ibid) は、イントネーションをプロソディ指導の一部として扱っているため、その効果は他のプロソディの要素の影響を受けている可能性がある。このように、既存の研究はイントネーション指導の有効性を示唆するものの、具体的な指導法に対する教育的示唆が十分でなく、またイントネーション指導の効果を取り出して検証することができていない。

## 3 研究・調査方法

### 3.1 研究課題

本研究では以下の三つの研究課題を設定する。

1. リズム・強勢とイントネーションの明示的指導を取り入れた音読活動は、学習者の発話の Comprehensibility を高めるか。
2. リズム・強勢とイントネーションの明示的指導を取り入れた音読活動は、リズム・強勢の明示的指導のみを取り入れた音読活動と比べ、学習者の発話の Comprehensibility の向上により効果的であるか。
3. リズム・強勢とイントネーションの明示的指導を取り入れた音読活動は、リズム・強勢の明示的指導のみを取り入れた音読活動と比べ、学習者の語彙の記憶の定着により効果的であるか。

研究課題1および2に答えるために実験1を実施し、また、研究課題3には実験2を実施した。

### 3.2 実験1

#### 3.2.1 実験の概要

実験1では、明示的なイントネーション指導が学習者の発話の Comprehensibility へ与える効果を検証するため、介入群と対照群を設けてそれぞれに指導を行い、事前テストおよび2種類の事後テストを用いて効果を検証した。指導は授業の一環として、5週間にわたり合計5回実施された。事前テストは初回指導時に、事後テストは全指導終了後の次の授業で実施した。なお、学習内容の公平性を確保するため、実験終了後の次の単元における音読活動の際に、後述のイントネーションの明示的指導を対照群に施した。

#### 3.2.2 実験参加者

本実験には、高等学校第1学年の少人数クラス4クラスが参加した。このうち2クラスが介入群、残りの2クラスが対照群に割り当てられた。研究参加にあたり、対象生徒には本研究の目的、データの取り扱いとプライバシー保護、また研究参加が任意であり、不参加や途中での取りやめによる不利益が一切生じないことを書面および口頭で説明した。その後、書面による研究同意を得た。指導介入の効果を正確に評

価するため、分析対象は、研究同意が得られた生徒のうち、実験期間中の全ての授業に出席した生徒に限定した。最終的に、介入群32名、対照群34名の計66名が本実験の分析対象となった。

### 3.2.3 指導手順

本実験における介入指導は、論理表現Ⅰの授業内で実施された。授業は全て著者が担当した。実験期間中の授業の単元目標は、自己紹介をテーマにスピーチ発表することであり、文法的な焦点は時制と助動詞に置かれた。授業では自己紹介スピーチのモデル文を使用し、音読を通じて構成と表現を学習することで、最終的に自己紹介スピーチができるようになることを目標とした。

モデル文は授業単元の目標に照らして作成された。スピーチの表現と文法事項を自然な形で導入するため、モデルスピーチは、1) オーディエンスへの挨拶、2) 中学生時代に努力したこと、3) 高校生活で取り組みたいこと、の3点で構成された。全5回の授業のため、それぞれ異なる内容のモデル文を5種類作成し、各授業で使用した(資料1)。モデル文の作成にあたって、生徒が親しみやすいトピックであること、および自らのスピーチに転用可能な表現を多く含むことに留意した。それぞれのモデル文の語数の平均は79語であった。

各群の指導の概要を表2に示す。また介入群の指導内容と Celce-Murcia et al. (2010) の枠組みとの対応を表3に示す。

■表2: 各群における指導概要

| 授業    | 介入群                                                                                                                                                                                          | 対照群                                                                                                                                                                                      |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1回目   | <ul style="list-style-type: none"> <li>● 事前テスト</li> <li>● 強勢・リズムの説明</li> <li>● イントネーションの説明</li> <li>● 教師によるモデルのリスニング</li> <li>● コーラスリーディング</li> </ul>                                        | <ul style="list-style-type: none"> <li>● 事前テスト</li> <li>● 強勢・リズムの説明</li> <li>● 教師によるモデルのリスニング</li> <li>● コーラスリーディング</li> </ul>                                                           |
| 2～5回目 | <ul style="list-style-type: none"> <li>● ペア音読</li> <li>● 表現・発音の確認</li> <li>● 教師によるモデルのリスニング</li> <li>● リズム・強勢・イントネーションの指導</li> <li>● コーラスリーディング</li> <li>● バズリーディング</li> <li>● 発表</li> </ul> | <ul style="list-style-type: none"> <li>● ペア音読</li> <li>● 表現・発音の確認</li> <li>● 教師によるモデルのリスニング</li> <li>● リズム・強勢・音素の発音指導</li> <li>● コーラスリーディング</li> <li>● バズリーディング</li> <li>● 発表</li> </ul> |
| 6回目   | <ul style="list-style-type: none"> <li>● 事後テスト1</li> <li>● 事後テスト2</li> </ul>                                                                                                                 | <ul style="list-style-type: none"> <li>● 事後テスト1</li> <li>● 事後テスト2</li> </ul>                                                                                                             |

■表3: 介入群の指導内容と Celce-Murcia et al. (2010) の枠組みとの対応

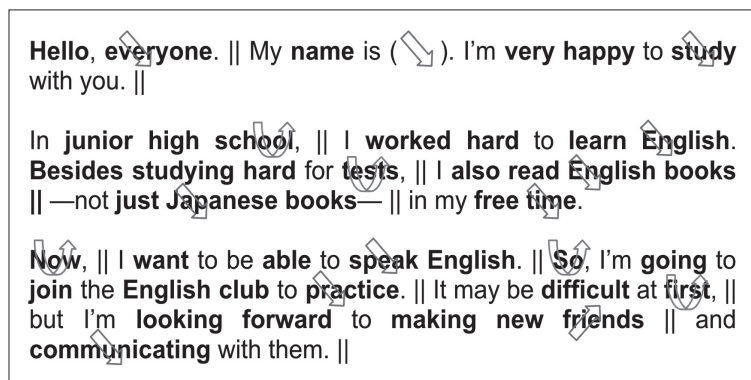
| 段階                          | 活動                                                | 留意点                  |
|-----------------------------|---------------------------------------------------|----------------------|
| 1. Description and Analysis | 強勢・リズム・イントネーションの明示的説明(初回)<br>ペア音読と表現・発音の確認(2～5回目) | 音声の意味・形式に対する意識の高揚    |
| 2. Listening Discrimination | 教師のモデル音読のリスニング発音方法の再検討(2～5回目)                     | 音声の形式に対する気づき         |
| 3. Controlled Practice      | 教師のモデルに続くコーラスリーディング                               | 音声の形式の習得             |
| 4. Guided Practice          | バズリーディング・発表                                       | 文章の意味と音声の形式の結びつけ     |
| 5. Communicative Practice   | 自己紹介スピーチ(単元のまとめ活動)                                | 意味重視の中で音声形式に対する意識の維持 |

初回授業では、まず事前テストを実施し、そのテストで用いられた文章を用いて音読指導を行った。最初に、介入群および対照群の双方に対し、強勢とリズムに関する明示的な指導を行った。具体的には、内容語と機能語の区別と、文強勢とリズムの等時性について説明した。続いて、対照群は教師による音読モデルを聴き、強勢とリズムを意識してコーラスリーディングを実施した。介入群に対しては、イントネーションに関する明示的な指導を行い、教師によるモデル音読を与え、リズム・強勢とイントネーションに意識を向けた中でコーラスリーディングを行った。

イントネーションの明示的説明では、Wells(2006)の3Tsのプロセス(Tonality, Tonicity, Tone)を基盤に、以下の内容を原則として指導した。

1. Tonality: 文章を一息で読める意味・文法のまとまりごとに区切る。
2. Tonicity: 意味のまとまりの内で最後の内容語においてピッチを変化させる。
3. Tone: ピッチの変化が文の途中で生じる場合は下降上昇調とする。並列表現の場合は下降上昇調ではなく上昇調とする。文の終わりである場合は下降調とする。
4. 対比や情報の新旧に関わり、強調すべきだと判断される情報が存在する場合は、原則2にかかわらず、当該箇所ではピッチを変化させる。

この原則に関する説明の後、図1のモデルを提示した。学習者はモデルを確認しながら教師のモデル音読を聴き、その後ピッチの変化に意識を向けながらコーラスリーディングを行った。



■図1:授業で提示されたイントネーションのモデル

第2回から第5回までの授業では、各回冒頭に介入群と対照群の両方がペア音読を行った。この活動では、まずモデル文を30秒間黙読し、その後1分以内にペアの相手に対して音読を行った。この活動は、学習者自身にアウトプットの機会を与えることで音声形式への気づきを促すこと、またその後の教師によるモデル音読のインプットが、自身の課題に関連づいたより効果的な学習になることを目的とした。

ペア音読後には、モデル文中の表現とペア音読時に誤読された単語や、生徒から質問のあった単語の発音を確認した。次に、両群ともにリズム・強勢への意識を高めるために、内容語に下線を引かせ、機能語との区別を促した。介入群にはこれに加え、意味のまとまりごとにスラッシュを引かせ(Tonality)、区切った句の中でピッチの変化が起これと思われる単語を選ばせ(Tonicity)、そのピッチの変化を文章中に記入させた(Tone)。これらの活動も、教師のモデル音読を聴く際に音声的特徴に意識を向けさせることを目的とした。

その後、教師のモデル音読を聴いた上で、介入群・対照群ともに内容語と機能語の区別を確認した。続いて、対照群は母音の挿入が起これやすい語頭の子音連結(山根, 2019)について発音指導を受けた。介入群には図1と同様の区切りとピッチ曲線のモデルを提示し、生徒が自身で記入したスラッシュとピッチ変化との比較を促した。その後、対照群はリズム・強勢と語頭の子音連結を、介入群はリズム・強勢とイントネーションを意識したコーラスリーディングをそれぞれ1回実施した。続いて、両群共に5分間のバズリー

ディングにより、個人練習の時間を確保した。最後に、介入群と対照群それぞれ数名の学習者に文章を音読させ、そのパフォーマンスに対してフィードバックを行った。6回目の授業が終了した後は、単元のまとめの活動として、自己紹介のスピーチ課題を行った。

### 3.2.4 事前テスト・事後テスト

事前テスト・事後テストともに、用意されたスクリプトを音読する課題を採用した。先行研究では即興で行う絵の描写課題 (Derwing et al., 1998; Zhang & Yuan, 2020) など、実際のコミュニケーション場面により近づけた課題を採用した実験もあるが、本研究の実験参加者の能力を考慮すると、語彙の選択や文法の正確性、または流暢さなど、より多くの変数が Comprehensibility に影響を持つことが想定され、イントネーションの指導の効果を正確に測定することが困難であると考え、読み上げの課題のみを実施することとした。

事前テストと事後テストに使用されたスクリプトは著者が作成した。実験期間中の授業で使用したモデル文と同様の構成で、語数の違いがテストの結果に影響を及ぼさないようにいずれも75語で作成した。

事後テストは、Saito and Saito (2017) と同様に、2種類実施した。事後テスト1は事前テストと同じスクリプトの音読課題で、既習の文章の音読における指導の効果測定のために実施した。事後テスト2は初見の文章で、未知の文脈における指導の効果測定のために実施した。

以下に事前・事後テストの手続きを説明する。実験参加者はスクリプトを渡され、準備時間として30秒間文章を黙読し、その後1分間で文章を読み切るように指示された。準備時間は読み方を検討する時間として、制限時間は過度に速度を落とした読み上げを防ぐために設定した。条件を揃えた状態でテストを実施するため、実験参加者は各自のデバイスで一斉に録音を行った。その際、音声の混在を防ぐため、各参加者はお互いに十分な距離をとった。

Comprehensibility の評価は Derwing, Munro and Wiebe (1998) と同様に、1~9の尺度で行われた (1=非常に理解が困難; 9=非常に理解が容易)。発話の評価者は英語母語話者1名と日本語母語話者1名の計2名で、双方共に応用言語学の修士号を持ち、日本の中等教育学校で英語教師として勤務している。評価者は Comprehensibility について、Derwing et al. (ibid) の定義をもとに書面と口頭で説明を受けた。評価の一貫性を向上させるため、評価者には1点、5点、9点のサンプル音声を与えられ、それらを聴いた上で評価に臨んだ。評価する順番がスコアに影響を与えることを避けるため、発話の音声ファイルは介入群と対照群を混ぜて提示した。

評価が終了した後、2名の評価者間の評定者信頼性 (Intraclass Correlation Coefficients: ICC) を評価するため、JASP (0.19.3) で級内相関係数を算出した。その結果、ICC (2,k) は0.876で95%信頼区間は0.836から0.907であった。この値は、良好な評定者間信頼性が示されたことを示唆している (Koo & Li, 2016)。最終的に二人の評価の平均を Comprehensibility のスコアとして分析することとした。

## 3.3 実験2

### 3.3.1 実験の概要

実験2は実験1の終了後に実施された。実験2では、音声面での改善ではなく、授業で扱われた語彙の記憶定着に対する、イントネーション指導を統合した音読活動の効果を検証した。実験は計2回の授業で実施され、実験1と同様に介入群と対照群を設けてそれぞれに異なる指導を実施し、事前・事後テストを用いて効果を測定した。事前テストは初回授業で実施され、その授業内で各群に対して指導が行われた。1週間後の授業で事後テストを実施し、それぞれの指導の効果を検証した。

### 3.3.2 実験参加者

実験1とそれぞれ同一の2クラスが介入群、2クラスが対照群として実験に参加した。研究同意が得られ

た参加者のうち、事前テストと事後テストの両方を受験した生徒を分析の対象とした。最終的に、介入群37名、対照群39名の計76名が本実験の分析対象となった。

3.3.3 実験手続き

本実験は、論理表現Ⅰの授業内で実施され、授業は著者が担当した。実験期間中における授業の単元目標は、学習者自身にとって思い入れのある場所についてプレゼンテーション発表することであり、実験はその単元の初回と2回目の授業で実施された。実験1と同様に、本授業ではモデル文が用いられ、音読を通じてその表現方法が指導された。当該モデル文は、後述する事前・事後テストで出題される英単語10語を含む形で著者により作成された(資料2)。

事前・事後テストにおいては、提示された英単語の意味を日本語で記述する形式の10問が出題された。英単語の選定にあたっては実験参加者の英語能力を考慮し、Oxford Learner's Dictionaries (Oxford University Press, n.d.)を参照の上、CEFR B2レベルおよびC1レベルの語彙から、名詞4語(perseverance; dedication; ambition; enthusiasm)、形容詞2語(sacred; located)、副詞1語(deeply)、動詞3語(reach; inspire; move)が選択された。事前テストと事後テストでは同一の英単語が出題され、問題の提示順序のみが変更された。テストは正解を1点、不正解を0点とし、10点満点による単純加算法で分析された。

各群の実験手続きの概要を表4に示す。

表4: 各群における実験手続きの概要

| 授業   | 介入群                                                                                                                                                                                                               | 対照群                                                                                                                                                                                                           |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 回目 | <ul style="list-style-type: none"><li>● 事前テスト</li><li>● 教師によるモデルプレゼンテーション</li><li>● 表現・発音の確認</li><li>● 教師によるモデルのリスニング</li><li>● リズム・強勢・イントネーションの指導</li><li>● コーラスリーディング</li><li>● バズリーディング</li><li>● 発表</li></ul> | <ul style="list-style-type: none"><li>● 事前テスト</li><li>● 教師によるモデルプレゼンテーション</li><li>● 表現・発音の確認</li><li>● 教師によるモデルのリスニング</li><li>● リズム・強勢・音素の発音指導</li><li>● コーラスリーディング</li><li>● バズリーディング</li><li>● 発表</li></ul> |
| 2 回目 | <ul style="list-style-type: none"><li>● 事後テスト</li></ul>                                                                                                                                                           | <ul style="list-style-type: none"><li>● 事後テスト</li></ul>                                                                                                                                                       |

実験参加者はまず事前テストを受験し、テスト後に出題された英単語の意味と発音を確認した。続いて、単元目標であるプレゼンテーションのイメージを掴むため、教師によるモデルプレゼンテーションを視聴した。その後モデル文が配布され、文中で使用されている表現および発音が確認された。次に、介入群と対照群の両方が内容語と機能語の区別を行った。介入群にはこれに加え、実験1と同様にTonality, Tonicity, Toneの原則に従い、モデル文に句の区切りとピッチの変化を書き込む作業を行った。続いて、教師のモデル音読を聴いた上で、介入群・対照群ともに内容語と機能語の区別を確認した。対照群には音素の発音指導を行い、介入群には区切り位置とピッチ曲線のモデルを提示し、生徒自身が書き込んだ内容との比較を促した。その後、両群ともにコーラスリーディングとバズリーディングで音読練習を行い、最後に数名の生徒が音読発表を行った。事前テストの問題用紙は、授業外での学習機会を排除するために回収し、事前テストの授業から1週間後の授業内で事後テストを実施した。

3.4 分析

3.4.1 研究課題1

それぞれの指導の効果を検証するために、群内における事前テストから事後テストへのスコアの推移を分析した。検定に先立ちデータの正規性を確認したところ、対照群の事後テスト1において正規性の前

提が満たされなかった( $p < .05$ )ため、対応のある $t$ 検定ではなく、ノンパラメトリック検定のウィルコクソンの符号付順位和検定を用いて分析を行った。表5は各群の事前テスト、事後テスト1・2の記述統計量を、表6はウィルコクソンの符号付順位和検定の結果を示す。

■表5: 事前テスト・事後テスト1・2記述統計量

| テストの種類 | 群名              | 平均値  | 標準偏差 | 中央値  | 四分位範囲 |
|--------|-----------------|------|------|------|-------|
| 事前テスト  | 対照群( $n = 34$ ) | 4.00 | 1.22 | 4.00 | 1.75  |
|        | 介入群( $n = 32$ ) | 3.72 | 1.51 | 3.50 | 1.50  |
| 事後テスト1 | 対照群( $n = 34$ ) | 4.78 | 1.40 | 4.50 | 1.50  |
|        | 介入群( $n = 32$ ) | 5.05 | 1.64 | 4.75 | 2.13  |
| 事後テスト2 | 対照群( $n = 34$ ) | 4.35 | 1.31 | 4.25 | 1.88  |
|        | 介入群( $n = 32$ ) | 4.78 | 1.78 | 4.50 | 2.00  |

■表6: ウィルコクソンの符号付順位和検定の結果

| 群名              | テストの種類 | $W$  | $z$   | $p$    | $r$   |
|-----------------|--------|------|-------|--------|-------|
| 対照群( $n = 34$ ) | 事後テスト1 | 5    | -4.33 | < .001 | -0.97 |
|                 | 事後テスト2 | 72   | -2.23 | .025   | -0.52 |
| 介入群( $n = 32$ ) | 事後テスト1 | 3    | -4.55 | < .001 | -0.99 |
|                 | 事後テスト2 | 21.5 | -4.13 | < .001 | -0.89 |

表6を見ると、事後テスト1において、両群ともに事前テストから有意にスコアを伸ばしていることがわかる。対照群の効果量は( $r = -0.97$ )、介入群の効果量は( $r = -0.99$ )であり、両群ともに指導によって大きな効果が得られたことが示唆されている(水本・竹内, 2010)。次に事後テスト2について、こちらも両群ともに事前テストから有意にスコアを伸ばしている。効果量は、対照群が( $r = -0.52$ )、介入群が( $r = -0.89$ )であり、両群において指導に大きな効果があったことが示唆された。

### 3.4.2 研究課題2

研究課題2では各群への指導の効果の比較をするために、群を独立変数、事前テストを共変量、事後テストを従属変数として、共分散分析を用いて分析した。共分散分析の実施にあたり、前提となる、1) 共変量と独立変数の独立性; 2) 従属変数の残差の正規性; 3) 残差の等分散性; 4) 共変量と従属変数の線形性; 5) 回帰直線の平行性の確認を行った。

1) 共変量と独立変数の独立性について、共変量である事前テストのスコアに各群間で有意差がないことを確認するため、対応のない $t$ 検定を実施した。その結果、両群の事前テストスコアに有意な差は認められなかった( $t(64) = 0.83$ ,  $p = .407$ )ため、共変量と独立変数が独立していることが確認された。次に2) 従属変数の残差の正規性に関して、Q-Qプロットを視覚確認した上で、シャピロ-ウィルク検定を実施した。その結果、事後テスト1・2いずれの残差も有意な逸脱がないことが認められた(事後テスト1  $W = 0.98$ ,  $p = .235$ ; 事後テスト2  $W = 0.97$ ,  $p = .164$ )。3) 残差の等分散性の確認のため、ルビン検定を実施した。その結果、事後テスト1・2のいずれにおいても分散は統計的に有意な差を示さなかった(事後テスト1  $F(1, 64) = 3.86$ ,  $p = .054$ ; 事後テスト2  $F(1, 64) = 0.71$ ,  $p = .402$ )ことから、等分散性が認められた。4) 共変量と従属変数の線形性は、事前テストと事後テスト1および2の散布図と回帰線から確認された。5) 回帰直線の平行性について、共変量と群の間の交互作用を確認したところ、事後テスト1において、この交互作

用は統計的に有意ではなく ( $F(1, 62) = 0.86, p = .357$ ), また事後テスト2においても同様に有意な交互作用は認められなかった ( $F(1, 62) = 0.26, p = .611$ )。これら1)から5)の結果から分析の前提は満たされたと判断し、共分散分析を実施した。表7は事後テスト1の共分散分析の、表8は事後テスト2の共分散分析の結果を、表9はそれぞれの調整後の平均値を示している。

■表7: 共分散分析の結果(事後テスト1)

| ケース   | 自由度 | 平均平方  | $F$    | $p$    | $\eta^2$ |
|-------|-----|-------|--------|--------|----------|
| 群     | 1   | 2.16  | 2.69   | .106   | 0.01     |
| 事後テスト | 1   | 97.68 | 122.10 | < .001 | 0.65     |

■表8: 共分散分析の結果(事後テスト2)

| ケース   | 自由度 | 平均平方  | $F$   | $p$    | $\eta^2$ |
|-------|-----|-------|-------|--------|----------|
| 群     | 1   | 7.43  | 7.42  | .008   | 0.05     |
| 事後テスト | 1   | 92.39 | 92.27 | < .001 | 0.57     |

■表9: 調整後の平均値(事後テスト1・2)

| テストの種類 | 群名               | 調整後平均値 | 95%信頼区間<br>(下限～上限) | 標準誤差 |
|--------|------------------|--------|--------------------|------|
| 事後テスト1 | 対照群 ( $n = 34$ ) | 4.64   | [4.34, 4.95]       | 0.15 |
|        | 介入群 ( $n = 32$ ) | 5.17   | [4.85, 5.49]       | 0.16 |
| 事後テスト2 | 対照群 ( $n = 34$ ) | 4.24   | [3.89, 4.59]       | 0.17 |
|        | 介入群 ( $n = 32$ ) | 4.91   | [4.56, 5.27]       | 0.18 |

事後テスト1・2のいずれにおいても、事前テストが事後テストの結果に対して有意な影響を与えていることが示された(事後テスト1,  $p < .001$ ; 事後テスト2,  $p < .001$ )。このことから、事前テストを共変量として調整したことは妥当であることが示された。

事後テスト1においては、群間で有意差は検出されず ( $F(1, 63) = 2.69, p = .106$ ), また効果量も小程度 ( $\eta^2 = 0.01$ ) であった(水本・竹内, 2010)。一方で、事後テスト2においては群間に有意差が認められ ( $F(1, 63) = 7.42, p = .008$ ), またその効果量は小から中程度 ( $\eta^2 = 0.05$ ) であった(水本・竹内, *ibid*)。これらの結果から、介入群と対照群の間には既習の文章の音読においては有意差がなかったものの、初見の文章においては介入群が有意に優れたパフォーマンスを示したことが示唆された。

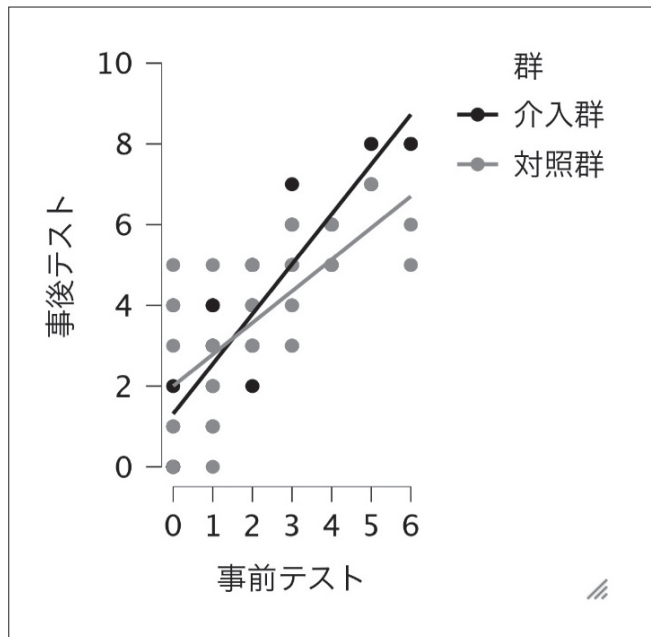
### 3.4.3 研究課題3

表10は各群の事前テスト、事後テストの記述統計量を示す。

■表10: 事前テスト・事後テスト記述統計量

| テストの種類 | 群名               | 平均値  | 標準偏差 | 中央値  |
|--------|------------------|------|------|------|
| 事前テスト  | 対照群 ( $n = 39$ ) | 2.26 | 1.60 | 2.00 |
|        | 介入群 ( $n = 37$ ) | 2.11 | 1.81 | 2.00 |
| 事後テスト  | 対照群 ( $n = 39$ ) | 3.77 | 1.81 | 4.00 |
|        | 介入群 ( $n = 37$ ) | 3.92 | 2.51 | 4.00 |

研究課題2と同様に、各群への指導の効果の比較をするために、共分散分析を用いて分析を試みた。検定に先立ち、共分散分析の前提の確認をしたところ、回帰直線の平行性において有意差が認められ( $F(1, 72) = 4.28, p < .001$ )、共変量と群の間に交互作用が存在することが示された。そのため図2の散布図と回帰直線を確認したところ、事前テストのスコアによって介入の効果に違いが生じる可能性が示唆された。具体的には、事前テストで低得点であった生徒に対しては対照群の指導が、高得点であった生徒に対しては介入群の指導がより効果的である可能性が示唆された。このことから、事前テストのスコアの分布を三分位に分割し(低得点群 [0-1点,  $n = 29$ ], 中得点群 [2-3点,  $n = 30$ ], 高得点群 [4-6点,  $n = 17$ ]), その効果を単純主効果分析でさらに詳細に分析することとした。



■図2: 散布図と回帰直線 (JASP 0.19.3で作成)

表11に三分位群別の事前テストおよび事後テストにおける記述統計量を示す。事後テストの平均値を確認すると、中得点群と高得点群では、介入群(中得点群  $M = 4.50$ ; 高得点群  $M = 6.89$ )が対照群スコア(中得点群  $M = 4.00$ ; 高得点群  $M = 5.88$ )を上回った。一方、低得点群では、介入群のスコア( $M = 1.81$ )が対照群( $M = 2.15$ )を下回った。

■表11: 三分位群別の事前テスト・事後テスト記述統計量

| 郡名                   | 事前テスト |      |      | 事後テスト |      |      |
|----------------------|-------|------|------|-------|------|------|
|                      | 平均値   | 標準偏差 | 中央値  | 平均値   | 標準偏差 | 中央値  |
| 介入群 低得点 ( $n = 16$ ) | 0.50  | 0.52 | 0.50 | 1.81  | 1.64 | 1.50 |
| 介入群 中得点 ( $n = 12$ ) | 2.25  | 0.45 | 2.00 | 4.50  | 1.38 | 5.00 |
| 介入群 高得点 ( $n = 9$ )  | 4.78  | 0.83 | 5.00 | 6.89  | 1.17 | 7.00 |
| 対照群 低得点 ( $n = 13$ ) | 0.54  | 0.52 | 1.00 | 2.15  | 1.73 | 2.00 |
| 対照群 中得点 ( $n = 18$ ) | 2.44  | 0.51 | 2.00 | 4.00  | 0.97 | 4.00 |
| 対照群 高得点 ( $n = 8$ )  | 4.63  | 0.92 | 4.00 | 5.88  | 0.64 | 6.00 |

表12は単純主効果分析の結果を示す。中得点群と高得点群の事後テストのスコアにおいて、介入群と対照群の間に有意差は認められなかった(中得点群  $F(1, 72) = 2.08, p = .154$ ; 高得点群  $F(1, 72) = 2.13, p = .149$ )。しかし、低得点群の事後テストのスコアでは、介入群と対照群の間に有意差が示され( $F(1, 72) = 4.75, p = .033$ )、介入群が対照群よりも有意に低いスコアであったことが明らかになった。

■表12: 単純主効果分析結果

| レベル  | 自由度 | 平均平方 | $F$  | $p$  |
|------|-----|------|------|------|
| 低得点群 | 1   | 7.29 | 4.75 | .033 |
| 中得点群 | 1   | 3.19 | 2.08 | .154 |
| 高得点群 | 1   | 3.27 | 2.13 | .149 |

## 4 結果および考察

### 4.1 研究課題1

音声面でのイントネーション指導の結果としては、介入群においては、2種類の事後テストそれぞれにおいて、事前テストから有意なスコアの上昇がみられた。また、効果量からは指導の効果が大きいことが示唆された。この結果は、音読活動を通じたリズムと強勢、そしてイントネーションの明示的指導が学習者の発話の Comprehensibility を効果的に向上させることを示している。これはプロソディ指導の効果を検証した先行研究 (Derwing et al., 1998; Ramirez Verdugo, 2006; Saito & Saito, 2017; Zhang & Yuan, 2020) の結果とも一致しており、本研究で考案した音読活動との統合的なイントネーションの指導が先行研究と同様に効果的であることが示唆された。

一方、対照群も同様に、2種類の事後テストにおいて有意なスコアの改善と大きな効果量が認められ、音読とリズム・強勢の明示的な指導のみでも高い効果が得られることが示唆された。この結果は、Isaacs & Trofimovich (2012) が Comprehensibility の評価に、リズム・強勢がとりわけ大きな影響を与えると指摘していることと一致しており、プロソディの構成要素の中でもリズム・強勢の指導が重要であることが改めて示されたと言える。また、Saito and Saito (2017) が指摘しているように、初級の日本人英語話者は母語の影響からそれぞれの音節を同じように強く読む傾向があり、強弱のリズムを作ることに困難がある。しかし、本指導を通してそういった習得課題を乗り越える効果が顕著に現れた可能性が考えられる。

ただし、本研究にはデザイン上の制約が存在する。具体的にはプロソディの指導を行わず、音読活動のみを行う統制群を設けていないため、音読活動自体の効果や、練習効果によるテスト形式への慣れの影響を排除することができていない。より厳密な効果検証が可能なデザインでの追試を行うことが今後の課題となる。

### 4.2 研究課題2

事後テスト1(既習の文章の音読課題)においては群間に有意差は認められなかった。しかし、事後テスト2(初見の文章の音読課題)では、介入群のスコアが対照群よりも有意に高いことが確認され、小から中程度の効果量が認められた。このことから、初見の文章の音読課題においては、リズム・強勢の指導に加えイントネーションの明示的指導を行うことがより効果的であったことが示された。

この結果は、それぞれの事後テストの性質と転移適切性処理 (Transfer Appropriate Processing: TAP) の観点から説明され得る。TAP によれば、学習時とテスト時の認知プロセスの一致度が高いほ

ど、学習した知識が転移しやすく、テストの際に学習効果が発揮されやすいとされている (Lightbown, 2008)。

事後テスト1で用いた文章では、練習時に教師のモデルが提示され、学習者はその模倣と反復練習を行った。そのため、事後テスト1ではそのモデルを記憶に基づいて再現することで、一定程度の Comprehensibility を達成することが可能だったと考えられる。この事後テストの課題の性質上、意味内容を十分に考えた上で発話するという認知プロセスは求められにくく、したがってイントネーションの指導の有無による効果の差は現れにくかったのではないかと考えられる。

一方、事後テスト2では、教師のモデルは与えられておらず、学習者自身が文章の意味内容を踏まえた上で、それを適切に表現することが求められた。この課題の性質は、文章の意味内容に基づき発音の方法を検討するという、介入群で行われた指導内容とより高い親和性を持っていたと考えられる。介入群の指導においては、学習者が意味内容や文法的な観点から文の区切り位置を判断し (Tonality)、その区切られた句の中で重要だと思われる語を選択する (Tonicity)、そしてその語に適切なピッチの変化を付与して発話する (Tone) という手続きが行われていた。この指導によって培われた、意味を解釈し発話するプロセスが、初見の文章を理解して読み上げるという事後テスト2の課題においてより効果的に転移したと解釈できる。

しかしながら、本研究には即興的な発話課題を事後テストに含んでいないため、実際のコミュニケーション場面における指導効果を十分に測定できていない。Derwing et al. (1998) や Zhang and Yuan (2020) が採用した即興型の発話課題を採用するなど、より実際のコミュニケーション場面に近い課題においても学習の転移が起きるかどうかを検証することが今後の重要な課題となる。

また、介入群の指導が学習者の発話に具体的にどのような変化をもたらし、その変化が Comprehensibility の向上にどう寄与したのかを、本研究では明確に特定できていない。Saito and Saito (2017) のように、プロソディの構成要素別に学習者の発話を詳細に分析することで、指導の効果をより具体的に把握することができるだろう。このような詳細な分析は、限られた授業時間の中で重点的に指導すべき項目を特定するための重要な知見をもたらすと考えられる。

### 4.3 研究課題3

事前テストで低得点であった層 (0-1点) の生徒に対しては、介入群の指導の効果が対照群よりも有意に低かったことが示された。この原因は、Type of Processing-Resource Allocation (TOPRA) モデルの観点から考えられる。TOPRA モデルによれば、学習者の認知資源は有限であり、複数の認知処理を行う際には認知資源が競合するため、ある処理に多く認知資源が割かれると、他の認知処理が阻害される可能性がある (Barcroft, 2002; Nakata, 2019)。

介入群の指導では、音調群、音調核、核音調の決定とその産出という処理が求められ、新出語彙の意味処理への認知資源が不足した可能性がある。介入群の中でも、低得点層の生徒にとっては文章中に未知語が多く、本来語彙学習に多くの認知資源を割く必要があったが、そこにイントネーションの産出処理が加わったことで認知資源が分散され、語彙学習が阻害された可能性が考えられる。この結果は、イントネーションの統合的な指導が、学習者の習熟度によっては、他の学習項目の習得を阻害する可能性を示唆している。

しかし、本研究における語彙知識の測定は、意味を記述する課題に限定されており、指導の効果の全体像を捉えきれていない。今後の課題として、語彙知識のより多角的な側面を対象に指導の効果を検証することが求められる。特に、TAP の観点からは、今回の指導が意味処理を伴う課題と親和性が高い可能性が示唆された。したがって、例えば、より実際のコミュニケーション場面に近い課題において学習した語彙の想起や使用が可能か検証することは、統合的な指導の有効性について重要な知見をもたらすと考えられる。

## 5 結論・教育的示唆

本研究は、イントネーション指導を音読活動に統合した指導法を考案し、その効果を検証することを目的として、高校1年生の日本人英語学習者を対象に実験を行った。実験1では全5回の指導を2群に分けて実施した。介入群ではイントネーションおよびリズム・強勢の明示的指導と音読活動を行い、対照群ではリズム・強勢の明示的指導と音読活動のみを行った。事前・事後テストでは文章の読み上げ課題を課し、その発話を Comprehensibility の観点から評価した。また、実験2では、同様の指導が語彙習得に与える影響を検証するため、未知語を含む文章を用いた音読活動を行い、学習した語彙の意味を記述するテストで効果を測定した。

本研究の結果から得られた主要な示唆は以下のとおりである。1) 音読活動を通じたリズムと強勢、イントネーションの明示的指導は学習者の発話の Comprehensibility の向上に有効である; 2) リズムと強勢の明示的指導と音読活動のみであっても、Comprehensibility の向上に高い効果が得られる; 3) 初見の文章の音読課題においては、イントネーションの明示的指導をリズム・強勢の指導に加えることが Comprehensibility の向上により効果的である; 4) イントネーションを意識した音読活動には一定の認知負荷を伴うため、学習者の習熟度によっては他の認知処理との間で認知資源の競合が起こる可能性がある。

これらの結果から、以下の教育的示唆が導かれる。まず、イントネーション指導の優先順位について、2) で示されたようにリズム・強勢の指導のみでも効果が期待できること、さらに英語のイントネーションがリズム・強勢と不可分であることを踏まえると、指導初期にはリズム・強勢の指導を中心にを行い、十分な習得が確認された段階でイントネーション指導へ移行することが妥当であると考えられる。次に、3) の示唆から、未知の文脈においても学習の転移が起こるよう、音読活動では教師のモデルの模倣や反復にとどまらず、文章の意味内容を踏まえて、学習者自身がイントネーションによる意味表現の方法を検討する時間を設けることが重要だと言える。また、4) で指摘された認知負荷について、学習者の習熟度や教材の難易度を慎重に吟味し、認知資源の競合が学習の阻害にならないよう指導計画を立てる必要がある。例えば、新出語彙が多く含まれる文章を扱う場合には、語彙理解と定着のための活動を十分に実施した上でイントネーションの指導に移行することが望ましい。

## 6 今後の展望

本研究は、先行研究で十分に記述されていなかったイントネーションの具体的な指導法について、教授可能性と学習可能性の観点から検討した。その上で、限られた時間で実行可能な統合的指導法を考案し、その効果を実証的に検証した点に大きな意義があると考ええる。この指導法が他の教室環境で応用され、さらなる教育的示唆が得られることで、指導法がより洗練されていくことが期待される。

今後の展望としては、考察で述べた事項に加え、以下の2点が挙げられる。第一に、本研究では教授可能性と学習可能性を考慮してイントネーションの指導内容を最小限に絞り込んだが、実際のコミュニケーションでは原則だけでは捉えきれない側面が存在する。そのため、インプットを通じて音声の特徴に対する学習者の気づきを促し、今回提示した原則に留まらないイントネーションに関する知識を増強する指導法を考案することが有益であると考ええる。例えば、Nego (2023) では、スピーチのリスニングを通じて音調核の配置と核音調の分析をすることで、話者のメッセージの解釈を促した。この実践例のように、具体的な教育実践を積み重ねることで、イントネーション指導の知見を深めることが重要であると考ええる。第二に、本研究では扱わなかった、Comprehensibility の向上に効果が高いと考えられるイントネーションの特徴を指導項目に追加することも今後の課題である。例えば、話題の転換を示すパラトーン (Paratone) は Comprehensibility に影響をあたえることが Wennerstrom (1998) の研究から明らかになっ

ている。このような特徴を指導項目に加え、その効果を検証することで、さらに効果的なイントネーション指導法の開発につながることが期待される。これらの研究課題に取り組むことで、より実践的かつ包括的なイントネーションの指導法の確立に貢献できるだろう。

## 謝辞

本研究の機会をくださいました公益財団法人 日本英語検定協会の皆様ならびに選考委員の先生方に、心より感謝申し上げます。特に、研究助言者の和泉伸一先生には丁寧かつ建設的なご助言を数多く賜りましたことを、ここに記して厚く御礼申し上げます。また本研究にご協力いただいた生徒の皆様の真摯な取り組みに深く感謝いたします。最後に、研究活動を支えてくれた家族に感謝の意を表します。なお、本報告書に関して開示すべき利益相反関連事項はありません。

## 引用文献

- Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, 52(2), 323-363. <https://doi.org/10.1111/0023-8333.00186>
- Celce-Murcia, M., Brinton, D., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268-294. <https://doi.org/10.1037//0096-3445.104.3.268>
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410. <https://doi.org/10.1111/0023-8333.00047>
- Gilbert, J. (2014). Intonation is hard to teach. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching*, (pp. 107-137). University of Michigan Press.
- Gordon, J., & Darcy, I. (2016). The development of comprehensible speech in L2 learners: A classroom study on the effects of short-term pronunciation instruction. *Journal of Second Language Pronunciation*, 2(1), 56-83. <https://doi.org/10.1075/jslp.2.1.03gor>
- Grant, L. (2014). Prologue to the myths: What teachers need to know. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching*, (pp. 1-33). University of Michigan Press.
- Isaacs, T., & Trofimovich, P. (2012). Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, 34(3), 475-505. <https://doi.org/10.1017/S0272263112000150>
- 門田修平 (2012). 「音読指導の実践 Q&A, Q1 音読にはどのような効果があるのでしょうか?」. 『英語教育』, 61(10), 10-12.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38(2), 301-315. <https://doi.org/10.1016/j.system.2010.01.005>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Levis, J. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.
- Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process*, (pp. 27-44). Multilingual Matters.
- Lindsey, G. A. (2019). *English after RP: Standard British pronunciation today*. Palgrave Macmillan.
- Low, E. L. (2006). A cross-varietal comparison of deaccenting and given information: Implications for international intelligibility and pronunciation teaching. *TESOL Quarterly*, 40(4), 739-761.
- 牧野武彦 (2005). 『日本人のための英語音声学レッスン』大修館書店.
- 水本篤, & 竹内理 (2010). 「効果量と検定力分析入門-統計的検定を正しく使うために」. 『外国語教育メディア学会(LET)関西支部 メソトロロジー研究部会2010年度報告論集』, 1, 44-70.
- 文部科学省 (2018). 『中学校学習指導要領(平成29年告示)解説 外国語編』開隆堂出版.
- Nakata, T. (2019). Learning words with flash cards and word cards. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies*, (pp. 304-319). Routledge.
- Nego, Y. (2024). Teaching prosodic features through reading aloud: A case in junior high school. *東京大学教育学部附属中等教育学校論集*, 67, 130-139.
- Nego, Y. (2025). Teaching English intonation: A case in junior high school. *東京大学大学院教育学研究科紀要*, 64, 189-193.
- O'Connor, J. D., & Arnold, G. F. (1973). *Intonation of colloquial English: A practical handbook* (2nd ed.). Longman.
- Oxford University Press. (n.d.). Oxford Learner's Dictionaries. Retrieved June 7, 2025, from <https://www.oxfordlearnersdictionaries.com/>

## 引用文献

- Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: Contemporary perspectives*. Palgrave Macmillan.
- Pickering, L. (2001). The role of tone choice in improving L2 communication in the classroom. *TESOL Quarterly*, 35(2), 233-255.
- Ramírez Verdugo, D. (2006). A study of intonation awareness and learning in non-native speakers of English. *Language Awareness*, 15(3), 141-159.
- Roach, P. J. (2009). *English phonetics and phonology: A practical course* (4th ed.). Cambridge University Press.
- Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. Continuum.
- 斎藤弘子 (2018). 「A longitudinal study of L2 English intonation—Does studying abroad make any difference?—」. In ことばを編む編集委員会 (Ed.), 『ことばを編む』 (pp. 27-37). 開拓社.
- Saito, Y., & Saito, K. (2017). Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Language Teaching Research*, 21(5), 589-608.
- 鈴木寿一, & 門田修平 (2012). 『英語音読指導ハンドブック: フォニックスからシャドーイングまで』 大修館書店.
- 鈴木祐一 (2024). 『あたらしい第二言語習得論: 英語指導の思い込みを変える』 研究社.
- 高橋愛紗 (2007). 「音声を併用したフレーズ・リーディングと音読が言語再生と保持に与える影響」. 『英語教育研究』, 30, 61-69.
- 竹内理 (2018). 「再考 音読活動」. 『Teaching English Now (TEN)』, 40, 5.
- 渡辺和幸 (1994). 『英語のリズム・イントネーションの指導』 大修館書店.
- Wells, J. C. (2006). *English intonation: An introduction*. Cambridge University Press.
- Wennerstrom, A. (1998). Intonation as cohesion in academic discourse: A study of Chinese speakers of English. *Studies in Second Language Acquisition*, 20(1), 1-25.
- 大和知史, & 磯田貴道 (2020). 「発音指導と他技能との統合: thought group の指導を通して」. 『神戸大学国際コミュニケーションセンター論集』, 16, 25-36.
- 山根繁 (2019). 『コミュニケーションのための英語音声学研究』 関西大学出版部.
- Zhang, R., & Yuan, Z. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition*, 42(4), 905-918. <https://doi.org/10.1017/S0272263120000121>

## 資料1：実験1音読指導用教材

## 事前テスト（初回授業教材）

Hello, everyone. My name is ( ). I'm very happy to study with you. In junior high school, I worked hard to learn English. Besides studying hard for tests, I also read English books — not just Japanese books — in my free time. Now, I want to be able to speak English. So, I'm going to join the English club to practice. It may be difficult at first, but I'm looking forward to making new friends and communicating with them.

## 2 回目授業教材

Hello, everyone. My name is ( ). It's a pleasure to be here with you. When I was a junior high school student, I spent a lot of time drawing pictures. Furthermore, I visited art museums to learn more about art. Now, my goal is to create better artwork. Therefore, I'll be taking an art class after school to improve my skills. It might make my life busier, but I'm eager to learn new techniques and exchange ideas with creative classmates.

## 3 回目授業教材

Hello, everyone. My name is ( ). I'm pleased to join you. Back in junior high school, I used to practice the piano at home every day. On top of that, I listened to many kinds of music to get a better understanding. At the moment, I'm working on expressing my own music. That's why I've decided to take part in a music workshop. It could be challenging at times. However, I hope to grow as a musician and collaborate with others.

## 4 回目授業教材

Hello, everyone. My name is ( ). I'm thrilled to be a member of this class. Since I entered this school, I have been interested in programming. I started by learning simple code online, and I enjoyed trying out small projects, like making games or websites. Currently, I'm trying to improve my programming skills. That's why I've decided to join a coding club outside of school. It might be a bit demanding at times, but I'm sure I'll be able to create something useful and work with people who share the same interest.

## 5 回目授業教材

Hello, everyone. My name is ( ). I'm delighted to be part of this class. As a junior high school student, I played tennis every morning. Also, I enjoyed practicing on my own to improve my techniques after school. These days, I am more interested in team sports rather than individual sports. That's why I've joined a soccer club and practice after school until evening. It is sometimes tiring, but I'm keen to improve my skills and to win games with great teammates.

## 事後テスト

Hello, everyone. My name is ( ). I'm pleased to be here with you. Back in junior high school, I helped my parents cook dinner every day. I also enjoyed making American dishes, like pizza and hamburgers. These days, I am more interested in cooking Japanese meals. That's why I've joined a cooking club and practice new skills after school. It's sometimes difficult, but I'll keep working hard with my friends and share delicious dishes with them.

Let me tell you about my sacred place, Koshien. Do you know where it is? Koshien is located in Hyogo Prefecture. It is famous for the national high school baseball tournament. For me, Koshien Stadium is more than just a place for baseball. It is a symbol of perseverance, dedication, and ambition. Players have made a great amount of effort to reach the place, and the game is filled with their enthusiasm. It always makes me feel inspired. Why don't you visit Koshien? You'll be deeply moved!

第37回 研究助成

**B** 実践部門・報告Ⅳ・英語能力向上をめざす教育実践

## 個別最適な学びを実現する 複線型の英語科授業を志向した探究的実践

研究代表者: 濱田 活仁 愛媛県／弓削商船高等専門学校 助教  
共同研究者: 新美 徳康 愛知県／豊田工業高等専門学校 講師(申請時: 広島県／広島大学大学院 在籍)  
福光 将仁 茨城県／筑波大学大学院 在籍

《研究助言者: 西垣 知佳子》

### 概要

本実践は、ライティングの「構想-執筆-修正」という一連のプロセスを経て意見文を書き上げる活動(プロセス・ライティング活動)において、学習者が各段階の学び方を選択する「複線型」の授業形態は、指導者が各段階の学び方を指定する「単線型」や「準複線型」の授業形態に比べて、学習者集団、学習者個人におけるライティングの質やライティング方略の変容にどのような違いをもたらすのかを把握することを目的とした。高専2年生3クラスに「単線型」、「準複線型」、「複線型」の授業形態を割り当て、2回の授業実践を行った。その結果、プロセス・ライティング活動による「複線型」授業は、「単線型」授業に比べると学習者集団におけるライティングの質の向上に一定の効果があることが示された。一方で、ライティング方略の使用の増大にはほとんど効果が示されなかった。また、「複線型」授業でライティングの質が向上していない学習者の一部には、単一の観点から単一の学び方を繰り返し選択して使用している特徴が発見された。本実践から、学習者が複数の観点から複数の学び方を選択して使用し、自身の学習スタイルを拡張していく意義が示唆された。

## 1 はじめに

### 1.1 実践の文脈

本論文の第一著者(筆者)は、地方の高等専門学校(以下、高専)において、高専2年生(高校2年生相当)を対象とした授業科目「総合英語」を担当している。「総合英語」では、「聞くこと」、「読むこと」、「書くこと」、「話すこと」の4技能を通して総合的な英語力を身に付けることと、学習者の自己調整を通じた自律性を育むことを授業目標に掲げている。

筆者は、令和6年度前期の授業において、診断的評価を行ったり、学習者がタスクに取り組む様子を観察したりする中で、対象校の学生が抱える課題の把握に努めた。そこで以下の2つの課題が浮かび上がった。1つ目の課題は、高専2年生である学生の多数が、中学校段階で身に付けるべき英語力の目標(特に、「書くこと」の目標)に達していないことが考えられる点である。筆者は、学生の「聞くこと」、「読むこと」、「書くこと」の英語力の実態を把握するために、令和6年度前期の第1回授業において、「令和5年度全国学力・学習状況調査(中学校英語)」調査問題を実施した。調査問題は、現行(平成29年告示)の中学校学習指導要領の目標に準拠した「聞くこと」、「読むこと」、「書くこと」の3領域を含む評価項目17問で構成されている。評価項目の難

易度は、Common European Framework of Reference for Languages(CEFR)レベルに換算するとA1からA2レベル相当が大多数を占めている(横浜国立大学, 2024)。「令和5年度全国学力・学習状況調査(中学校英語)」調査問題の全国平均(当時の中学3年生)と対象校平均(高専2年生)の正答率は、表1に示す通りである。

■表1: 「令和5年度全国学力・学習状況調査(中学校英語)」調査問題の平均正答率

|                | 聞くこと  | 読むこと  | 書くこと  | 3領域   |
|----------------|-------|-------|-------|-------|
| 全国平均(当時の中学3年生) | 58.9% | 51.7% | 24.1% | 46.1% |
| 対象校(高専2年生)     | 68.1% | 59.1% | 28.4% | 53.2% |

注. 全国平均(当時の中学3年生)の数値は、文部科学省・国立教育政策研究所(2023, p. 10)を参照した。

対象校平均は、全国平均に比べると、「聞くこと」で9.2ポイント、「読むこと」で7.4ポイント、「書くこと」で4.3ポイント上回った。しかしながら、調査問題が中学校の目標に準拠していることを踏まえると、当時の中学3年生が抱えていた課題と同様に、対象校の学生の英語力が中学校の目標に到達しているとは言い難い。特に「書くこと」の正答率は、対象校平均と全国平均ともに3割に満たなかった。この診断結果から、対象校の学生は、当時の中学3年生と同様に、「書くこと」で大きな課題を有していると考えられた。

2つ目の課題は、英語学習において機械翻訳や生成AIといったデジタルツールを適切に扱えていないと考えられる学生が散見される点である。対象校では、英語科を含む一般科目の授業においてタブレットやPCといった1人1台端末を活用するための教室環境が未だ完備されておらず、代替手段として全ての学生が所持している私用スマートフォンを介してデジタルツールを活用する機会を設定している。例えば、前期の授業内において、まとまりのある説明文を書くタスクに取り組ませた際には、半数以上の学生は英語の文章を自分の力で書き進める中で、英語表現が分からないときには、自身のスマートフォンで機械翻訳を使用し、辞書のように参照しながら書き進める様子が観察された。自分の力でライティングを遂行していくことを前提としながらも、自分の力だけでは書き進めることができない部分においてデジタルツールを援用している。こうした学生は、自分の力を向上させるためにデジタルツールを適切に使用しており、自律的に学習を進めることができていると筆者は考える。その一方で、一定数の学生は、英語の文章を自分の力で書くプロセスを一度も踏むことなく、機械翻訳や生成AIが出力した英語の文章をそのまま全て引き写す、いわゆるコピー＆ペーストをしてタスクを終えている様子が散見された。また、生成AIが産出したモデル文を参照しながら自身のライティング・テストの解答を修正する活動を高専学生に実施した際には、生成AIのモデル文を活用する妥当な理由のもとで適切に修正されている箇所(例えば、ライティング・タスクの目的・場面・状況等に照らして適切な内容を選択しようと修正がなされた箇所など)が59.1%確認された。一方で、適切な修正であるものの妥当な理由が確認できない修正箇所(例えば、語数を満たすために生成AIが産出した表現を盲信的に引き写して修正がなされた箇所など)が31.5%を占めており、一部の学生は生成AIのモデル文を無批判に引き写している実態が確認されている(濱田他, 2025)。こうした学生は、自分の力を向上させるためにデジタルツールを適切に活用しているとは言い難く、英語力の向上を目的として自律的に学習を進めることができていないことが考えられた。

このように令和6年度前期の授業を通して、対象校の学生は、まとまりのある英語の文章を書く力と、自律的に英語の文章を書くために必要なデジタルツールの活用を含むライティング方略の使用に関して、大きな課題を抱えていることが明らかになった。これらの課題を把握した筆者は、前期の授業では、話すこと(発表)のパフォーマンス課題に向けた準備段階としてライティング活動を位置付けていたが、後期の授業ではライティング・タスク、具体的には意見文を書くプロセス・ライティング活動を重点的に取り扱うことを着想した。本科目「総合英語」では、「聞くこと」、「読むこと」、「書くこと」、「話すこと」の4技能を通した総合的な英語力の育成を目指していることから、後期において「書くこと」に特化した授業のみを展開することは現実的ではない。ただ一方で、教科書の各レッスンの単元末においてライティング・タスクを一貫して設定することは、実現可能な手立てであると筆者は考えた。例えば、各レッスンに関わ

るトピックについて「主張-理由-具体-結論」といった構成を意識した意見文を書くライティング・タスクを繰り返し行うことは、中学生程度の英語力を有する学生には適した学び直しの機会となり、ライティングの質の向上が期待できる。また、「構想-執筆-修正」といったライティングのプロセスにおいて、学習者がデジタルツールをはじめとした複数の学び方を選択し、学習過程を自己調整していくプロセス・ライティング活動は、自律的に英語の文章を書くために必要なライティング方略の使用の増大にも期待できる。このように、令和6年度前期の授業における総合的な省察を踏まえて、後期ではプロセス・ライティング活動の授業実践を通して、対象校が抱えているライティングの質やライティング方略における課題を解決するためのアクション・リサーチを展開していくことを筆者は構想した。

しかしながら、共同研究チームにおけるインサイダーである実践者(第一著者)とアウトサイダーである共同研究者(第二著者と第三著者)により研究計画を具体化する協議では、アクション・リサーチの展開に際して踏むべきステップとして、第一に対象校に適したプロセス・ライティング活動の授業デザインを探究する実践研究の必要性が浮かび上がった。プロセス・ライティング活動では、「構想-執筆-修正」といった各段階を踏みながら学びを進めることが考えられるが、その全ての段階で指導者が指定した「同一」の学び方で学習を進める授業形態、いわゆる「単線型」の授業形態がバリエーションの一つとして考えられる。例えば、「構想」と「執筆」の段階では自分の力のみで行うように指定し、「修正」の段階において機械翻訳を使って確認をしながら学習を進めるような授業形態は、学習者による自己調整の余地は限られてくるが、対象校の学生には馴染みのある学び方である。他方で、「構想-執筆-修正」の全ての段階で学習者が「複数」の学び方から選択して学習を進める授業形態、いわゆる「複線型」の授業形態もまたバリエーションの一つとして考えられる。例えば、「構想-執筆-修正」の全ての段階において、教科書や授業資料を参照、辞書や機械翻訳で確認、生成AIに相談、他の学習者に相談など、学習者個人が複数の学び方から選択しながら学習を進めるような授業形態は、自己調整の余地は大きい。一方で、濱田他(2025)で確認されたように、生成AIの産出文を無批判に全て引き写すような不適切な学び方を選択する学習者への懸念がある。これら「単線型」や「複線型」、もしくはこれらの折衷案となるような「準複線型」といった授業形態の違いは、学習者集団、学習者個人双方のライティングの質やライティング方略の変容にどのような違いをもたらすのだろうか。こうした問いに答えるための研究を行うことが、アクション・リサーチに向けたステップとして必要ではないだろうか。つまり、対象校に適した授業形態を明らかにするために、より良いプロセス・ライティング活動の授業デザインを探究する実践研究を行う必要性を我々は認識した。

そこで本実践では、意見文を書くプロセス・ライティング活動で、「構想-執筆-修正」の各段階の学び方を学習者が選択する「複線型」の授業形態が、「単線型」や「準複線型」に比べて、学習者集団、学習者個人におけるライティングの質やライティング方略の変容にどのような違いをもたらすのかを把握することを目的とした。

## 1.2 「個別最適な学び」を実現する授業形態

実践での授業形態を検討するにあたり、我々はまず、我が国の教育政策として掲げられている「個別最適な学び」の概念に着目した。初等中等教育では現在、「個別最適な学び」と「協働的な学び」の一体的な充実を図ることを目指し、デジタル学習基盤である1人1台端末を活用した授業展開が推進されている(文部科学省, 2021)。そのうち「個別最適な学び」は学習者の視点から整理された概念であり、教師の視点から「個に応じた指導」とも言い換えられるが、「指導の個別化」、「学習の個性化」の2種類に整理されている。「指導の個別化」とは、「一定の目標を全ての児童生徒が達成することを目指し、個々の児童生徒に応じて異なる方法等で学習を進めること」(pp.7-8)と定義されている。一方で、「学習の個性化」とは、「個々の児童生徒の興味・関心等に応じた異なる目標に向けて、学習を深め、広げること」(p.8)を意味している。こうした「個別最適な学び」を英語の授業に当てはめて検討した津久井(2024)は、前者の「指導の個別化」について「学習者の特性や学習進度、実態に応じて、教師が学習環境をデザインすること」との解釈を示し、「英語教師は単元全体のdesignerとして、全ての生徒が一定の目標に到達することを目的に、3年間・1年

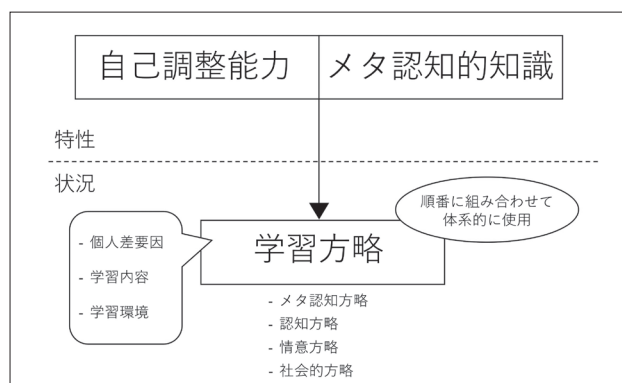
間の到達目標を俯瞰し、既習事項や題材の系統性を考慮して単元のゴールとなる言語活動を構想する力が今まで以上に求められる」(p.1)と指摘している。本研究では、この「指導の個別化」の視座を高専に組み込み、英語科における「書くこと」の資質・能力として設定した到達目標を全ての学生が達成することを目指し、プロセス・ライティング活動を通して個々の学生に応じて異なる学び方を選択し、その学び方を自己調整しながら学習を進める「個別最適な学び」を志向した授業デザインを構想する。

また「個別最適な学び」を実現する具体的な授業デザインの1つとして、高橋(2022)は「複線型」という授業形態を提唱している。「複線型」授業とは、各個人のペースで学習過程を自己決定しながら学びを進める授業形態を指す。他方、一斉指導のように全員が同一方法や同一のペースで学習活動を進める従来型の授業形態は「単線型」と形容される。初等中等教育で担保されている1人1台端末が整備されたICT環境は、前者の「複線型」授業における学習過程をより豊かにする役割を担うと主張している。

ただ一方で、高等教育機関に位置付けられる高専では、全ての学校で学習用端末を活用した英語科授業が展開できているとは言い難い。高専におけるICT環境の整備は各学校に対応が委ねられており、経済的事情を抱えた学生が一定数在籍している対象校では、学生に対して一律に学習用端末の持参を求めることが難しい。こうした実情を踏まえた対応として、ほぼ全ての学生が所持している私用スマートフォンを学習用端末の代替として用いることが考えられる。例えば、ライティング活動では、学習者が書いた英文を自ら修正する際に、教科書や授業資料を参照しながら修正していく学び方に加えて、スマートフォンを介して機械翻訳や生成AIを使いながら修正していく学び方も選択することができる。したがって本研究では、学習用端末が整備されていない環境下であっても、学習者が自身の適性に応じて学び方を選択し、学習過程を自己調整していく「複線型」授業を目指す。

### 1.3 外国語学習における学習方略と学習スタイル

我が国における教育政策用語として「個別最適な学び」の概念が整理された一方で、英語教育学や第二言語習得(Second Language Acquisition: SLA)の学術領域では、個人差に対応した外国語学習や指導において重要となる学習者要因の概念が整理されている。Li et al.(2022)は、学習者要因の分類として、認知的(cognitive)差異、意欲的(conative)差異、情動的(affective)差異、社会文化的(sociocultural)・人口能動的(demographic)差異の4つに整理している。学習者要因の4つの分類の中には、さらに具体的な要因が多様に存在しているが、本研究との関連性が強く考えられる概念は、認知的(cognitive)差異に分類される「学習方略」と「学習スタイル」である。そのうち、学習方略は、「学習者が意識的に使用する学習方法・行動」を意味する(竹内, 2014, p.236)。図1は、Mizumoto(2018)をもとに学習方略と関連する概念との関係を整理したものである。



■図1: 学習方略と関連する概念との関係(Mizumoto, 2018をもとに筆者が作成)

学習方略の選択と使用は、自己調整能力とメタ認知的知識によって引き起こされる思考や行動と解釈

される。学習者は、個人差要因や学習内容、学習環境に応じて、複数の方略から選択し、順番や組み合わせを工夫しながら体系的に使用している (Mizumoto, 2018)。

また、学習者が選択する方略や具体的な学習方法には好みがあり、そうした個人の特性は「学習スタイル」と呼ばれるが、単一の学習スタイルに固執するのではなく、複数の学習方法を組み合わせて学習スタイルを拡張していくことが、英語学習においては意義があると考えられている (鈴木, 2024)。本研究が目指す「複線型」授業では、学習者が複数の学び方から主体的に選択する機会を設けることから、学習スタイルの拡張の有無によりライティングの質や学習方略の変容に違いが生じることが予想される。

## 1.4 自己調整学習に向けたライティング方略

外国語としての英語 (English as a Foreign Language: EFL) でのライティングにおいて発揮される学習方略は、心理的尺度を用いた質問紙を使って測定が可能とされている (Teng & Zhang, 2016)。本研究ではプロセス・ライティング活動の実践において学習者が学び方の選択を自己調整する機会を設定することから、Teng and Zhang (2016) により開発された質問紙にて測定可能な「自己調整学習に向けたライティング方略 (writing strategies for self-regulated learning)」の分類を取り扱う。表2は、自己調整に向けたライティング方略を構成する4つの上位方略と9つの下位方略を示している。

■表2: 自己調整学習に向けたライティング方略の分類 (Teng & Zhang, 2016をもとに筆者が作成)

| 上位方略                                | 下位方略                                                                                              |
|-------------------------------------|---------------------------------------------------------------------------------------------------|
| 認知 (Cognition) 方略                   | テキスト処理 (Text Processing) 方略<br>授業記憶 (Course Memory) 方略                                            |
| メタ認知 (Metacognition) 方略             | 内容構想 (Idea Planning) 方略<br>目標志向 (Goal-Oriented Monitoring and Evaluating) 方略                      |
| 社会的行動 (Social Behavior) 方略          | ピア学習 (Peer Learning) 方略<br>フィードバック操作 (Feedback Handling) 方略                                       |
| 動機づけ調整 (Motivational Regulation) 方略 | 興味喚起 (Interest Enhancement) 方略<br>自己対話 (Motivational Self-Talk) 方略<br>感情制御 (Emotional Control) 方略 |

Teng and Zhang (2016) は、自己調整学習に向けたライティング方略 (writing strategies for self-regulated learning) を測定するために、認知 (Cognition) 方略、メタ認知 (Metacognition) 方略、社会的行動 (Social Behavior) 方略、動機づけ調整 (Motivational Regulation) 方略の4つの上位方略と9つの下位方略に基づいた40項目の質問紙を開発した。大学生790名の回答データから確証的因子分析 (CFA) と構造方程式モデリング (SEM) を用いて3種の理論モデルを比較した結果、下位方略に基づく9因子モデルが良好な適合度を示し、一つの上位因子である自己調整 (Self-Regulation) によって説明可能であることが確認された。また、重回帰分析により下位方略がライティング能力を予測するかどうかを検証したところ、9因子のうちテキスト処理 (Text Processing) 方略、内容構想 (Idea Planning) 方略、目標志向 (Goal-Oriented Monitoring and Evaluating) 方略、フィードバック操作 (Feedback Handling) 方略、自己対話 (Motivational Self-Talk) 方略、感情制御 (Emotional Control) 方略の6因子が、学生のEFLライティング能力の有意な予測因子であることが示された。一方で、授業記憶 (Course Memory) 方略、ピア学習 (Peer Learning) 方略、興味喚起 (Interest Enhancement) 方略の3つの因子は、有意な予測因子ではないことが確認された (Teng & Zhang, 2016)。

本研究では、機械翻訳や生成AIなどのデジタルツールを活用したフィードバックを学習者が受容して

自身のライティングに取り入れることも、社会的行動(Social Behavior)方略におけるフィードバック操作(Feedback Handling)方略の一種として捉えている。しかしながら、Teng and Zhang (2016)が開発した質問紙には、現代的なAI技術に関する項目は含まれておらず、AI技術を扱ったフィードバック操作に関する方略は測定できない。そのため本研究では、Teng and Zhang (2016)が開発した質問紙をもとに一部の項目を再構成した質問紙を新たに作成し、ライティング方略を測定することにした。このようにプロセス・ライティング活動の実践を通してライティングの質の変容とライティング方略の変容を検証することは、本研究の授業デザインを評価する上で重要な指標の1つであると我々は考えた。

## 1.5 本研究の研究課題

本研究の目的は、意見文を書くプロセス・ライティング活動で、構想、執筆、修正の各段階の学び方を学習者が選択する「複線型」の授業形態は、「単線型」や「準複線型」に比べて、学習者集団、学習者個人におけるライティングの質やライティング方略の変容にどのような違いをもたらすのかを把握することである。これら目的を達成するために明らかにすべき研究課題(RQs)として、以下の2点を設定した。

RQ1

プロセス・ライティング活動における「複線型」という授業形態は、「単線型」や「準複線型」に比べて、学習者集団のライティングの質やライティング方略の変容に有効なのか。

RQ2

「複線型」授業では、学習者集団の学び方の選択と使用にどのような傾向が見られ、また学習者個人の学び方の選択と使用には、学習者間でどのような違いが見られるのか。

## 2 実践方法

### 2.1 参加者

参加者は、対象校に通う高専2年生、学科の異なる3クラス116名である。参加者の英語習熟度は、診断的テストとして実施した「令和5年度全国学力・学習状況調査(中学校英語)」調査問題の結果から、CEFRレベルA1からA2レベル相当と推定された。また、クラス単位での英語ライティングの質を、後述する事前テストの結果から比較すると、3クラスのうち2クラスは同程度の水準であったが、1クラスは他の2クラスに比べてやや低い水準であった。このことから、授業形態の設定に際しては、英語ライティングの質が同程度の水準を示した2クラスに対して、それぞれ「単線型」と「複線型」の授業形態を割り当て、両方でライティングの質の変容の違いを比較できる設計とした。一方で、事前テストで英語ライティングの質がやや低かった1クラスには「準複線型」の授業形態を割り当て、分析における他クラスとの比較は参考程度に留めることにした。

なお、参加者に対しては、実践者より本研究の目的と方法、データの取り扱い、プライバシーの保護等の説明が丁寧に行われ、データ提供の可否について書面にて同意を確認した。参加者のうち、本研究へのデータ提供に同意し、全4回の対象授業に全て参加した76名を分析の対象とした。

### 2.2 対象授業

2024年10月から12月にかけて行われた週1回90分の科目「総合英語」の後期授業(全15回)の一部、全4回を研究の対象授業とした。高等学校における「英語コミュニケーションⅡ」に準じた本科目では、高等学校検定教科書「VISTA English Communication II」(三省堂)を使用している。1つのLessonごとに(a)教科書を扱っ

た活動(主に「聞くこと」,「読むこと」), (b)ライティング・タスク(主に「書くこと」), (c)発表・やり取り活動(主に「話すこと」)を3週の授業で順に行う単元構成とし, Lesson 5からLesson 8までの4セットを実施した。そのうち本研究では, (b)ライティング・タスクに焦点を当て, 後期授業の第3回に事前テスト, 第6回に授業実践1回目, 第9回に授業実践2回目, 第12回に事後テスト, 計4回のライティング・タスクに取り組んだ。

## 2.3 ライティング・タスク

計4回のライティング・タスクでは, いずれも社会的な話題に関する問いに対して50語程度で賛否を述べる構成(主張-理由-具体-結論)の意見文に取り組んだ。各ライティング・タスクで取り上げた問いは, 表3に示す通りである。

■表3: ライティング・タスクで取り上げた問い

| 授業回             | 問い                                                                                                                          |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------|
| 第3回<br>〈事前テスト〉  | Should we limit the growth of foreign tourists?<br>(外国人観光客の増加を制限すべきか?)                                                      |
| 第6回<br>【実践1回目】  | Should we promote the replacement of work by robots and AI?<br>(ロボットやAIによる業務代替を促進すべきか?)                                     |
| 第9回<br>【実践2回目】  | Should we move from school-based club activities to community-based club activities?<br>(学校単位での部活動から地域単位でのクラブ活動に移行していくべきか?) |
| 第12回<br>〈事後テスト〉 | Should we stop keeping animals in zoos?<br>(動物園で動物を飼育することをやめるべきか?)                                                          |

それぞれ前時に取り扱った教科書Lessonのトピックに関連した問いを提示し, ライティング・タスクに取り組む直前に, 問いに関連した日本語のニュース動画を全員で視聴し, 背景知識を補足した。

## 2.4 授業実践

文章を書き上げるプロセスは, ①書く前の構想の段階(planning), ②実際に書いている執筆の段階(translating), ③書き上げた文章を修正する段階(revising)の3段階に大きく分けられる(保田, 2024)。

本実践で展開するプロセス・ライティング活動でも, ①構想(約10分), ②執筆(約20分), ③修正(約20分)の3つの段階に分け, それぞれの段階で費やす時間の目安を示し, 段階ごとに分割したワークシートを共通して使用することにした(資料1, 2)。

■表4: プロセス・ライティング活動における授業形態の分類

| 授業形態     | 具体的な学習過程                                                                                                              |
|----------|-----------------------------------------------------------------------------------------------------------------------|
| (A) 単線型  | ①構想と②執筆において自分の力のみで意見文を書き, ③修正では書きたかった英語表現を機械翻訳で確認して修正する授業形態                                                           |
| (B) 準複線型 | ①構想と②執筆においては自分の力のみで意見文を書くが, ③修正では学習者個人が, 教科書や授業資料を参照, 辞書機能や機械翻訳で確認, 生成AIに相談, 他の学習者に相談など, 複数の学び方を自己調整して自身の意見文を修正する授業形態 |
| (C) 複線型  | ①構想, ②執筆, ③修正, 全ての段階において学習者個人が複数の学び方を自己調整する授業形態                                                                       |

授業形態は表4が示すように、クラスごとに(A)単線型、(B)準複線型、(C)複線型を割り当てた。(A)単線型では、自分の力で構想と執筆を行い、機械翻訳を使って修正をするという、前期授業で対象校の多くの学生が行っていた学び方を、全ての学生に指定した。(B)準複線型では③修正の段階、(C)複線型では①構想、②執筆、③修正、全ての段階において、学習者個人が、教科書や授業資料を参照、辞書機能や機械翻訳で確認、生成AIに相談、他の学習者に相談など、複数の学び方を自己調整して自身の意見文を修正する授業形態とした。これらは、複数の学習方略や複数の学習方法を選択し、順番や組み合わせを工夫しながら体系的に使用するという、先行研究における学習方略(Mizumoto, 2018)や学習スタイルの拡張(鈴木, 2024)の考え方を取り入れた授業形態である。なお、いずれの授業形態の実践においても、指導者は学習者個人へのフィードバックは実施せず、学び方の選択肢に関する質問への助言に留め、次時に全体へのモデル文の例示と解説を実施した。

また、プロセス・ライティング活動の、(B)準複線型では③修正の段階、(C)複線型では①構想、②執筆、③修正、全ての段階において、表5が示す「学び方」の選択肢を指導者から学習者に対して提示した。学び方の選択肢は、令和6年度前期の授業においてライティング・タスクを実施した際に、対象校の学生が実際にやっている様子が観察された学び方を取り上げた。学び方への共通理解を図るために、プロセス・ライティング活動に取り掛かる前に、それぞれの学び方の説明と例示を指導者が行った。なお、学び方の選択肢は該当するワークシートの下部に明示されており、学習者は選択した学び方に費やした時間を大まかな分単位で記録した(資料1, 2)。

■表5: 学習者が自己調整する学び方の選択肢

|   | 学び方の選択肢               | 略称                 |
|---|-----------------------|--------------------|
| a | 自分の力のみ                | (a) 自力             |
| b | 内容・構成について教科書・授業資料を参照  | (b) 内容面(教科書・授業資料)  |
| c | 内容・構成についてインターネット検索で確認 | (c) 内容面(ネット検索)     |
| d | 内容・構成について生成 AI に相談    | (d) 内容面(生成 AI)     |
| e | 内容・構成について他の学生に相談      | (e) 内容面(他の学生)      |
| f | 英語表現について教科書・授業資料を参照   | (f) 言語面(教科書・授業資料)  |
| g | 英語表現について辞書機能・機械翻訳で確認  | (g) 言語面(辞書機能・機械翻訳) |
| h | 英語表現について生成 AI に相談     | (h) 言語面(生成 AI)     |
| i | 英語表現について他の学生に相談       | (i) 言語面(他の学生)      |
| j | その他                   | (j) その他            |

2.5

ライティング評価ルーブリック

ライティング・プロダクトを評価するルーブリックは、岩本(2021)が作成したルーブリックを使用した(資料3)。本ルーブリックは、国際バカロレア科目「English B」のライティング評価ルーブリックを中学生1年生にも理解できる表現に訳して作成されたものである。【内容面】と【言語面】の2観点から構成されており、【内容面】は主に論理的な主張ができているかどうかを、【言語面】は主に文法や語彙の使用が適切かどうかを、それぞれ0点から5点までの尺度で評価する。

本研究における実践の文脈として、学習者が中学生程度のライティング習熟度を有していること、また授業においてルーブリックを使用した経験が乏しいことを踏まえて、岩本(2021)のルーブリックを使用することが適当であると判断した。また、本ルーブリックは指導者から学習者に共有され、「自分の力のみ」

によるライティングが求められる事後テストにおいて、2観点ともに3点以上になることを到達目標に掲げている旨を、授業のたびに繰り返し説明がなされた。

### 3 分析データ

研究課題(RQs)に応えるために、(RQ1)プロセス・ライティング活動における「複線型」という授業形態は、「単線型」や「準複線型」に比べて、学習者集団のライティングの質やライティング方略の変容に有効なのに対しては、(1a)事前・事後テストとして実施したライティング・タスクの採点データ、(1b)ライティング方略を測定する質問紙への回答データを分析の対象とした。また、(RQ2)「複線型」授業では、学習者集団の学び方の選択と使用にどのような傾向が見られ、また学習者個人の学び方の選択と使用には、学習者間でどのような違いが見られるのかに対しては、(1a)と(1b)のデータに加えて、(2a)「複線型」授業において学生が使ったワークシートで得られた学び方の選択肢と時間の記録データ、(2b)その学び方を選択した理由を記述した追加質問紙の回答データを分析の対象とした。

(1a)事前・事後テストとして実施したライティング・タスクの採点データは、資料3で示したライティング評価ルーブリック(岩本, 2021)を用いて0点から5点の【内容面】、【言語面】の2観点、計10点満点の【総得点】で採点したデータである。ライティングの採点は、第一著者、第二著者、第三著者の3名が評価者となり、うち2名ずつが分担して行った。評価者2名の間で採点した得点に分かれたライティングは、もう1名の評価者が得点を判断して確定させた。なお、評価者3名は採点に取り組む前に、一部の対象授業の欠席により分析対象外のデータとなったライティング・タスク10名分を用いて、事前の採点トレーニング、採点基準についての協議を行い、共通認識を図った。(1b)ライティング方略を測定する質問紙は、Teng and Zhang(2016)が開発した質問紙をもとに30項目を作成した(資料4)。事前・事後テストのそれぞれの後に、参加者が質問紙に回答する時間を設定し、各項目について7段階のリッカート尺度(1.当てはまらない/2.ほとんど当てはまらない/3.あまり当てはまらない/4.どちらとも言えない/5.やや当てはまる/6.かなり当てはまる/7.非常に当てはまる)で尋ねた。

(2a)「複線型」授業において学生が選択した学び方の選択肢と時間の記録データは、ワークシート下部の欄に学生が記入した記録を使用した。プロセス・ライティング活動における①構想、②執筆、③修正それぞれの段階で使用するワークシート及び記録欄を3つ設け、どの段階でどの学び方を選択し、大まかな分単位でどのくらいの時間を費やしたのか、学生は自分で記録しながらプロセス・ライティング活動を進めた(資料1, 2)。(2b)事後テストの次時(第13回授業)において、「複線型」授業を割り当てたクラスのみ、実践2回目の中で学生が選択した学び方とその理由を尋ねる追加質問紙を行い、記述回答データを収集した。

本研究では、116名の参加者のうちデータ提供に同意し、対象授業4回の全てに参加し、(1a)と(1b)のデータセットが不足なく対応していた76名を分析の対象データとした。授業形態ごとの対象データとなる参加者の内訳は、(A)単線型が25名、(B)準複線型が27名、(C)複線型が24名となった。(2a)と(2b)のデータは、誤った記録によりデータが欠損していた2名を除外し22名を分析の対象とした。

### 4 結果と発見

#### 4.1 授業形態の違いによる「ライティングの質」の変容の違い

はじめに、「単線型」、「準複線型」、「複線型」という授業形態別に「ライティングの質」の変容を検証する。表6に事前と事後のライティング・テストの総得点及び観点別得点の記述統計を、図2に学習者個人の変容を可視化した散布図を授業形態別に示す。

表6: 事前と事後ライティング得点の記述統計

|     | A. 単線型 (n = 25) |      |      |      |      | B. 準複線型 (n = 27) |      |      |      |      | C. 複線型 (n = 24) |      |      |      |      |
|-----|-----------------|------|------|------|------|------------------|------|------|------|------|-----------------|------|------|------|------|
|     | 事前              |      | 事後   |      | d    | 事前               |      | 事後   |      | d    | 事前              |      | 事後   |      | d    |
|     | M               | SD   | M    | SD   |      | M                | SD   | M    | SD   |      | M               | SD   | M    | SD   |      |
| 内容面 | 2.16            | 1.37 | 2.32 | 0.99 | 0.15 | 1.48             | 0.70 | 1.93 | 0.78 | 0.69 | 2.00            | 1.18 | 2.50 | 1.32 | 0.43 |
| 言語面 | 2.00            | 1.04 | 2.24 | 0.93 | 0.24 | 1.41             | 0.69 | 1.78 | 0.80 | 0.50 | 2.00            | 0.98 | 2.33 | 1.09 | 0.32 |
| 総得点 | 4.16            | 2.36 | 4.56 | 1.89 | 0.20 | 2.89             | 1.37 | 3.70 | 1.49 | 0.64 | 4.00            | 2.09 | 4.83 | 2.33 | 0.41 |

注1. 総得点は、内容面(0点～5点)、言語面(0点～5点)の2観点、計10点満点である。  
注2. 効果量dの大きさの目安は、d = 0.2で小程度、d = 0.5で中程度、d = 0.8で大程度の効果量を示す(水本・竹内, 2008)。

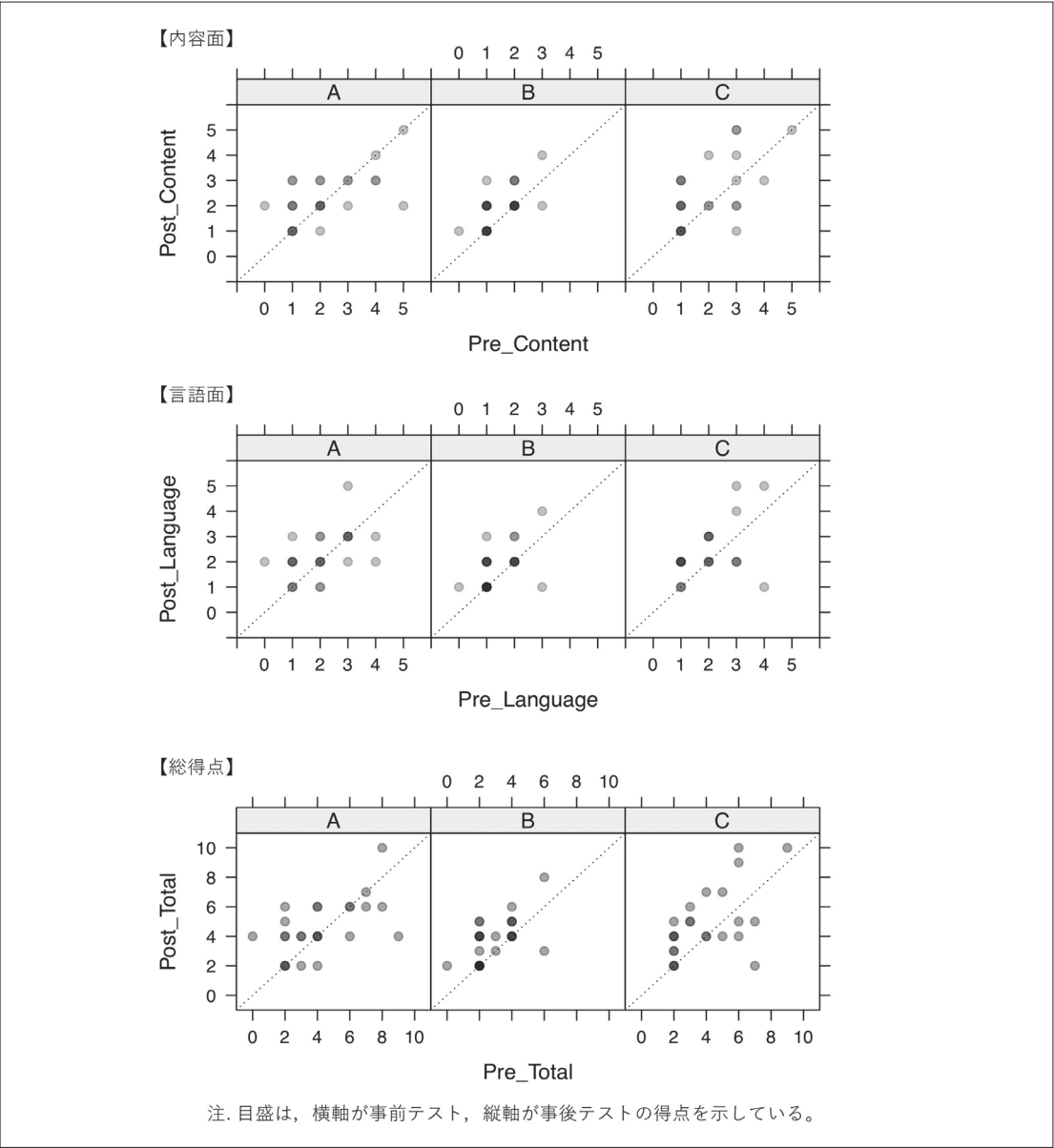


図2: 事前と事後のライティング得点の散布図

まず、事前テストにおける【総得点】の平均値を見ると、(A)単線型が4.16、(B)準複線型が2.89、(C)複線型が4.00を示しており、事前のライティングの質は(A)と(C)が同程度の水準と判断できる。一方で、(B)はやや低い水準であることが確認された。よって以降、授業形態の違いによる変容の違いを比較する際には、(A)単線型と(C)複線型を中心に扱い、(B)準複線型は参考に留めることにする。

授業形態の違いによる学習者集団におけるライティングの質の変容を比較するために、表6の事前・事後テストの平均値の差を標準化した効果量 $d$ に着目する。(A)と(C)を比較すると、(C)のほうが総得点及び観点別の効果量が大きいことから、プロセス・ライティング活動における「複線型」授業には、単線型授業と比較して一定の効果が示唆されたものと解釈できる。特に、【内容面】の効果量を見ると、(A)単線型は $d = 0.15$ で、小程度の効果量が確認されたのに対し、(C)複線型は $d = 0.43$ で、中程度の効果量が確認されたのは特徴的な発見である。

授業形態の違いによる学習者個人におけるライティングの質の変容を比較するために、図2の散布図におけるプロット位置に着目する。散布図の点線より左上に位置するプロットは、事前から事後テストで各得点が〈上昇〉した学習者の群を示している。〈上昇〉した人数の割合を比較すると、【総得点】では(A)単線型が10名(40.0%)と半数を下回ったのに対して、(C)複線型が14名(58.3%)と半数を上回り、(A)に比べて(C)の方が高い割合を示した。【内容面】や【言語面】といった観点別に見ても、【総得点】と同じく(A)に比べて(C)の方が高い割合であった。なお(B)準複線型では、【総得点】が〈上昇〉したのが14名(51.9%)となり、(C)複線型に準じた割合を示した。散布図の点線に位置するプロットは、事前から事後テストで各得点が〈不変〉であった学習者の群を示している。〈不変〉であった人数の割合を比較すると、【総得点】では(A)単線型が9名(36.0%)に対して(C)複線型が5名(20.8%)となり、(A)に比べて(C)の方が低い割合を示した。【内容面】や【言語面】においても、同様の傾向であった。なお(B)準複線型では、【総得点】が〈不変〉であった群が12名(44.4%)と、半数近い割合を示し、【内容面】では14名(51.9%)、【言語面】では15名(55.6%)と、半数を超える値を示した。散布図の点線より右下に位置するプロットは、事前から事後テストで各得点が〈下降〉した学習者の群を示している。〈下降〉した人数の割合を比較すると、【総得点】では(A)単線型が6名(24.0%)に対して(C)複線型が5名(20.8%)となり、(A)と(C)で同程度の割合を示した。【内容面】や【言語面】においても、同様の傾向であった。なお(B)準複線型では、総得点及び各観点項目のいずれにおいても〈下降〉した群は1名(3.7%)に留まった。

これらの結果から、「単線型」授業に比べると、「複線型」授業はライティングの質の向上に一定の効果があったことが確認された。一方で、個別最適な学びを志向した「複線型」授業であっても、全ての学習者のライティングの質を向上させるには及ばなかった。これは「複線型」授業において学習者個人が選択した学び方が、得点の伸びに寄与した学習者(上昇群)と、得点の伸びに寄与するに及ばなかった学習者(不変群)と、得点に負の影響を与えた学習者(下降群)がそれぞれ存在していたことを示唆している。

## 4.2 授業形態の違いによる「ライティング方略」の変容の違い

次に、プロセス・ライティング活動における「単線型」、「準複線型」、「複線型」という授業形態の違いによる「ライティング方略」の変容の違いを把握するために、7段階のリッカート尺度による事前・事後質問紙への回答データを用いて、4つの上位方略ごとの平均値を尺度得点とした記述統計を表7に、9つの下位方略ごとの平均値を尺度得点とした記述統計を表8にそれぞれ示す。

表7及び表8より、「複線型」授業では【認知】方略に位置づくテキスト処理方略の使用のみが増大傾向を示し、その他の下位方略の使用は全て減少傾向を示すことが明らかになった。また「単線型」授業に比べると、「複線型」授業ではテキスト処理方略の使用を増大する効果が小さかった。そして、「単線型」授業では増大傾向を示した目標志向方略といった【メタ認知】方略、興味喚起方略や感情制御方略といった【動機づけ調整】方略は、いずれも「複線型」授業では減少傾向を示すという変容の違いが明らかになった。

■表7: 事前と事後ライティング方略の記述統計（上位方略別）

| 上位<br>方略   | A. 単線型 ( $n = 25$ ) |      |      |      |       | B. 準複線型 ( $n = 27$ ) |      |      |      |       | C. 複線型 ( $n = 24$ ) |      |      |      |       |
|------------|---------------------|------|------|------|-------|----------------------|------|------|------|-------|---------------------|------|------|------|-------|
|            | 事前                  |      | 事後   |      | $d$   | 事前                   |      | 事後   |      | $d$   | 事前                  |      | 事後   |      | $d$   |
|            | $M$                 | $SD$ | $M$  | $SD$ |       | $M$                  | $SD$ | $M$  | $SD$ |       | $M$                 | $SD$ | $M$  | $SD$ |       |
| 認知         | 4.20                | 1.07 | 4.38 | 1.06 | 0.27  | 4.44                 | 0.89 | 4.49 | 0.95 | 0.05  | 4.33                | 1.08 | 4.30 | 1.21 | -0.04 |
| メタ<br>認知   | 4.46                | 1.02 | 4.43 | 1.26 | -0.03 | 4.43                 | 0.90 | 4.46 | 1.00 | 0.03  | 4.75                | 1.21 | 4.40 | 1.16 | -0.21 |
| 社会的<br>行動  | 4.75                | 1.19 | 3.92 | 1.52 | -0.73 | 5.28                 | 0.77 | 4.18 | 1.41 | -0.75 | 5.07                | 1.19 | 3.74 | 1.42 | -0.87 |
| 動機づけ<br>調整 | 4.40                | 0.95 | 4.60 | 1.52 | 0.22  | 4.60                 | 0.93 | 4.49 | 0.92 | -0.15 | 5.07                | 1.15 | 3.82 | 1.27 | -0.81 |

注. 効果量  $d$  の大きさの目安は,  $d = 0.2$  で小程度,  $d = 0.5$  で中程度,  $d = 0.8$  で大程度の効果量を示す (水本・竹内, 2008)。

■表8: 事前と事後ライティング方略の記述統計（下位方略別）

| 下位<br>方略          | A. 単線型 ( $n = 25$ ) |      |      |      |       | B. 準複線型 ( $n = 27$ ) |      |      |      |       | C. 複線型 ( $n = 24$ ) |      |      |      |       |
|-------------------|---------------------|------|------|------|-------|----------------------|------|------|------|-------|---------------------|------|------|------|-------|
|                   | 事前                  |      | 事後   |      | $d$   | 事前                   |      | 事後   |      | $d$   | 事前                  |      | 事後   |      | $d$   |
|                   | $M$                 | $SD$ | $M$  | $SD$ |       | $M$                  | $SD$ | $M$  | $SD$ |       | $M$                 | $SD$ | $M$  | $SD$ |       |
| テキスト<br>処理        | 4.31                | 1.24 | 4.73 | 1.26 | 0.55  | 4.67                 | 0.92 | 4.91 | 1.04 | 0.24  | 4.51                | 1.16 | 4.70 | 1.15 | 0.22  |
| 授業<br>記憶          | 4.09                | 1.15 | 4.04 | 1.38 | -0.05 | 4.21                 | 1.13 | 4.07 | 1.24 | -0.12 | 4.15                | 1.57 | 3.90 | 1.56 | -0.17 |
| 内容<br>構想          | 4.76                | 1.33 | 4.43 | 1.63 | -0.22 | 4.49                 | 1.28 | 4.48 | 1.43 | -0.01 | 4.94                | 1.30 | 4.50 | 1.80 | -0.29 |
| 目標<br>志向          | 4.16                | 1.26 | 4.44 | 1.25 | 0.24  | 4.36                 | 0.89 | 4.43 | 0.89 | 0.09  | 4.56                | 1.61 | 4.31 | 1.16 | -0.15 |
| ピア<br>学習          | 4.65                | 1.46 | 3.13 | 1.83 | -0.81 | 4.96                 | 1.04 | 3.74 | 1.69 | -0.72 | 4.64                | 1.49 | 2.90 | 1.78 | -0.92 |
| フィード<br>バック<br>操作 | 4.85                | 1.36 | 4.71 | 1.59 | -0.10 | 5.60                 | 0.82 | 4.62 | 1.40 | -0.62 | 5.50                | 1.10 | 4.58 | 1.66 | -0.56 |
| 興味<br>喚起          | 3.95                | 1.20 | 4.20 | 1.36 | 0.21  | 4.17                 | 1.25 | 4.14 | 1.18 | -0.03 | 4.04                | 1.33 | 3.07 | 1.70 | -0.81 |
| 自己<br>対話          | 4.73                | 1.44 | 4.85 | 1.32 | 0.08  | 4.88                 | 0.96 | 4.86 | 1.06 | -0.01 | 4.65                | 1.45 | 4.42 | 1.56 | -0.25 |
| 感情<br>制御          | 4.53                | 1.16 | 4.75 | 1.30 | 0.16  | 4.77                 | 1.31 | 4.48 | 1.26 | -0.23 | 4.35                | 1.27 | 3.99 | 1.49 | -0.25 |

注. 効果量  $d$  の大きさの目安は,  $d = 0.2$  で小程度,  $d = 0.5$  で中程度,  $d = 0.8$  で大程度の効果量を示す (水本・竹内, 2008)。

■表9: 「単線型」と「複線型」とで変容に特徴的な違いを示したライティング方略項目の一部

| 項目                                    | A. 単線型 ( $n = 25$ ) |      |      |      |       | C. 複線型 ( $n = 24$ ) |      |      |      |       |
|---------------------------------------|---------------------|------|------|------|-------|---------------------|------|------|------|-------|
|                                       | 事前                  |      | 事後   |      | $d$   | 事前                  |      | 事後   |      | $d$   |
|                                       | $M$                 | $SD$ | $M$  | $SD$ |       | $M$                 | $SD$ | $M$  | $SD$ |       |
| 2. 文法の間違いをチェックする。                     | 4.32                | 1.70 | 4.84 | 1.34 | 0.40  | 5.04                | 1.40 | 4.88 | 1.45 | -0.09 |
| 4. 論理的な一貫性があるか構成をチェックする。              | 4.48                | 1.81 | 4.44 | 1.64 | -0.03 | 4.13                | 2.03 | 4.54 | 1.59 | 0.26  |
| 14. 英作文の授業で, 知識やスキルの習得度を自己評価する。       | 4.20                | 1.41 | 4.52 | 1.50 | 0.18  | 4.96                | 1.81 | 4.46 | 1.25 | -0.22 |
| 21. AIからのアドバイスに基づいて, 自分の英作文を改善しようとする。 | 4.44                | 1.96 | 5.00 | 1.89 | 0.24  | 5.88                | 1.15 | 4.88 | 2.05 | -0.48 |
| 29. 英作文をあきらめなくなったとき, 「書き続けよう」と意識する。   | 4.36                | 1.60 | 5.00 | 1.66 | 0.33  | 4.29                | 1.71 | 4.38 | 1.76 | 0.06  |

注. 効果量  $d$  の大きさの目安は,  $d = 0.2$  で小程度,  $d = 0.5$  で中程度,  $d = 0.8$  で大程度の効果量を示す (水本・竹内, 2008)。

さらに具体的なライティング方略の変容の違いを把握するために、「単線型」と「複線型」とで変容に対照的な違いが見られた質問項目を一部抽出し、表9に示す。

例えば、テキスト処理方略に含まれる項目「2.文法の間違いをチェックする」方略の効果量を見ると、(A)単線型は $d = 0.40$ で上昇に中程度の効果量を、(C)複線型は $d = -0.09$ でほとんど不変の効果量を示した。一方で、同じテキスト処理方略に含まれる項目「4.論理的な一貫性があるか構成をチェックする」方略の効果量を見ると、(A)単線型は $d = -0.03$ でほとんど不変の効果量を、(C)複線型は $d = 0.26$ で上昇に小程度の効果量を示した。これらの結果は、同じテキスト処理方略であっても、「単線型」と「複線型」とでは変容に効果を示した具体的な方略が異なっていたことを示唆している。全ての学習者が機械翻訳を使って主にライティングの言語面への修正を施していた「単線型」授業では、学習者の「2.文法の間違いをチェックする」方略(言語面)の使用は多かったものの、「3.論理的な一貫性があるか構成をチェックする」方略(内容面)の使用は少なかった可能性が考えられる。一方で、複数の学び方から選択してライティングの内容面への修正も可能であった「複線型」授業では、「3.論理的な一貫性があるか構成をチェックする」方略(内容面)の使用が多かった可能性が推察される。なお、表9に示したその他の項目(14, 21, 29)の方略使用においては、いずれも「単線型」授業では使用が増大する傾向を示し、「複線型」授業では不変もしくは使用が減少する傾向を示した。

これらの結果から、「単線型」授業で使用が増大する傾向が確認されたライティング方略であっても、「複線型」授業では使用が減少したり不変であったりする傾向が明らかとなった。「単線型」授業では、学習者が使用する単一の学び方を指導者が指定していたことから、その単一の学び方に付随したライティング方略を全ての学習者が使用することになり、特定のライティング方略の使用に集中しやすい環境にあった可能性が考えられる。一方の「複線型」授業では、学習者が複数の学び方から選択して使用していたことから、それら複数の学び方に付随して使用するライティング方略も分散し、特定のライティング方略の使用には集中しなかった可能性が考えられる。

### 4.3 「複線型」授業における学習者集団の学び方の選択と使用の傾向

続いて、「複線型」授業における学習者集団の学び方の選択と使用にどのような傾向が見られるのかを把握するために、「複線型」授業において学習者が使用したワークシートの記録データを用いて、実践1回目と実践2回目における「学び方の選択肢」の項目別の【選択人数】の集計を表10に示す。また観点等別における【使用時間】の記述統計を表11に、【選択個数】の記述統計を表12にそれぞれ示す。

表10より実践1回目から2回目への「学び方の選択肢」の項目別の【選択人数】の変化に着目すると、②執筆する段階で(a)自力を部分的にも選択した学生が、実践1回目では19名(86.4%)に対し、2回目では14名(63.6%)と減少していた。一方で、同じく②執筆する段階から(h)言語面(生成AI)を選択した学生が、実践1回目では3名(13.6%)に対し、2回目では8名(36.4%)に増加していた。これは修正に入る前の執筆の段階から、すでに自力ではなく生成AIを選択する傾向が増したことを意味しており、「単線型」や「準複線型」と違い、必ずしも自力での執筆を強いていない「複線型」授業だからこそ見られた特徴的な発見である。表11の結果を概観すると、実践1回目と2回目とで各段階で観点等別の学び方に費やす【使用時間】には大きな変化は確認されなかった。つまり、本実践における学習者集団の学び方の選択と使用の傾向として、①構想する段階では、概ね10分ほど費やす中で、主に「自力」で構想しており、②執筆する段階では、概ね15分ほど費やす中で、主に「言語面」に関する学び方に時間を取りながらも部分的に「自力」でも執筆しており、③修正する段階では、概ね20分ほど費やす中で、主に「言語面」に関する学び方に大部分の時間を割いて修正している傾向が明らかとなった。

さらに、表12より実践1回目から2回目への「学び方の選択肢」の観点別の【選択個数】の変化に着目すると、構想、執筆、修正いずれの段階の「上記9項目(a)～(i)合計」の欄を見ても、選択個数の平均値が減少している。この結果から、実践1回目から2回目にかけて学習者が学び方を精選したことが考えられるが、「複線型」授業が学習スタイルの拡張に寄与するという筆者らの予想に反して意外な発見となった。

■表10: 学習者が使用し記録した「学び方の選択肢」の項目別の【選択人数】の集計 (n = 22)

| 段落  | 学び方の選択肢(項目別)       | 実践1回目 |       |      | 実践2回目 |       |      |
|-----|--------------------|-------|-------|------|-------|-------|------|
|     |                    | 選択人数  | 選択率   | 平均時間 | 選択人数  | 選択率   | 平均時間 |
| ①構想 | (a) 自力             | 21    | 95.5% | 8分   | 22    | 100%  | 9分   |
|     | (b) 内容面(教科書・授業資料)  | 1     | 4.5%  | 2分   | 1     | 4.5%  | 2分   |
|     | (c) 内容面(ネット検索)     | 5     | 22.7% | 5分   | 3     | 13.6% | 4分   |
|     | (d) 内容面(生成AI)      | 1     | 4.5%  | 15分  | 0     | —     | —    |
|     | (e) 内容面(他の学生)      | 0     | —     | —    | 0     | —     | —    |
|     | (f) 言語面(教科書・授業資料)  | 1     | 4.5%  | 2分   | 0     | —     | —    |
|     | (g) 言語面(辞書機能・機械翻訳) | 3     | 13.6% | 7分   | 4     | 18.2% | 4分   |
|     | (h) 言語面(生成AI)      | 0     | —     | —    | 0     | —     | —    |
|     | (i) 言語面(他の学生)      | 0     | —     | —    | 0     | —     | —    |
|     | (J) その他(休憩など)      | 3     | 13.6% | 7分   | 0     | —     | —    |
| ②執筆 | (a) 自力             | 19    | 86.4% | 6分   | 14    | 63.6% | 7分   |
|     | (b) 内容面(教科書・授業資料)  | 3     | 13.6% | 5分   | 0     | —     | —    |
|     | (c) 内容面(ネット検索)     | 5     | 22.7% | 6分   | 3     | 13.6% | 8分   |
|     | (d) 内容面(生成AI)      | 1     | 4.5%  | 10分  | 3     | 13.6% | 7分   |
|     | (e) 内容面(他の学生)      | 1     | 4.5%  | 2分   | 0     | —     | —    |
|     | (f) 言語面(教科書・授業資料)  | 2     | 9.1%  | 3分   | 1     | 4.5%  | 10分  |
|     | (g) 言語面(辞書機能・機械翻訳) | 12    | 54.5% | 8分   | 11    | 50.0% | 7分   |
|     | (h) 言語面(生成AI)      | 3     | 13.6% | 9分   | 8     | 36.4% | 8分   |
|     | (i) 言語面(他の学生)      | 1     | 4.5%  | 1分   | 1     | 4.5%  | 1分   |
|     | (J) その他(休憩など)      | 2     | 9.1%  | 4分   | 0     | —     | —    |
| ③修正 | (a) 自力             | 13    | 59.1% | 5分   | 11    | 50.0% | 6分   |
|     | (b) 内容面(教科書・授業資料)  | 0     | —     | —    | 2     | 9.1%  | 4分   |
|     | (c) 内容面(ネット検索)     | 4     | 18.2% | 7分   | 5     | 22.7% | 9分   |
|     | (d) 内容面(生成AI)      | 4     | 18.2% | 10分  | 2     | 9.1%  | 11分  |
|     | (e) 内容面(他の学生)      | 1     | 4.5%  | 5分   | 0     | —     | —    |
|     | (f) 言語面(教科書・授業資料)  | 2     | 9.1%  | 2分   | 1     | 4.5%  | 2分   |
|     | (g) 言語面(辞書機能・機械翻訳) | 13    | 59.1% | 14分  | 12    | 54.5% | 14分  |
|     | (h) 言語面(生成AI)      | 7     | 31.8% | 11分  | 8     | 36.4% | 11分  |
|     | (i) 言語面(他の学生)      | 3     | 13.6% | 8分   | 1     | 4.5%  | 10分  |
|     | (J) その他(休憩など)      | 3     | 13.6% | 9分   | 2     | 9.1%  | 10分  |

注. 選択人数の平均時間は、項目別の選択人数を分母とし、選択していない者(0分)は除外している。

■表11: 「学び方の選択肢（その他を除く）」の観点等別における【使用時間】の記述統計（ $n = 22$ ）

| 段落                      | 学び方の選択肢（観点等別）        | 実践1回目    |            |            | 実践2回目    |            |            |
|-------------------------|----------------------|----------|------------|------------|----------|------------|------------|
|                         |                      | <i>M</i> | <i>Mdn</i> | <i>Max</i> | <i>M</i> | <i>Mdn</i> | <i>Max</i> |
| ①構想                     | 自力 (a)               | 8.00     | 7.00       | 21.00      | 8.91     | 9.00       | 16.00      |
|                         | 内容面 (b), (c),(d),(e) | 1.64     | 0.00       | 15.00      | 0.68     | 0.00       | 5.00       |
|                         | 言語面 (f),(g),(h),(i)  | 0.91     | 0.00       | 8.00       | 0.68     | 0.00       | 5.00       |
|                         | 上記9項目 (a)～(i) 合計     | 10.55    | 10.00      | 24.00      | 10.27    | 10.00      | 16.00      |
| ②執筆                     | 自力 (a)               | 5.14     | 4.50       | 15.00      | 4.64     | 5.00       | 15.00      |
|                         | 内容面 (b), (c),(d),(e) | 2.68     | 0.00       | 13.00      | 2.14     | 0.00       | 13.00      |
|                         | 言語面 (f),(g),(h),(i)  | 5.95     | 6.00       | 18.00      | 6.77     | 7.00       | 15.00      |
|                         | 上記9項目 (a)～(i) 合計     | 13.77    | 15.00      | 22.00      | 13.55    | 13.00      | 24.00      |
| ③修正                     | 自力 (a)               | 2.82     | 1.50       | 12.00      | 2.82     | 1.00       | 10.00      |
|                         | 内容面 (b), (c),(d),(e) | 3.23     | 0.00       | 29.00      | 3.32     | 0.00       | 20.00      |
|                         | 言語面 (f),(g),(h),(i)  | 12.77    | 15.50      | 25.00      | 12.36    | 11.00      | 25.00      |
|                         | 上記9項目 (a)～(i) 合計     | 18.82    | 18.50      | 34.00      | 18.50    | 17.50      | 31.00      |
| 3段階における上記9項目 (a)～(i) 総計 |                      | 43.14    | 45.00      | 57.00      | 42.32    | 42.50      | 60.00      |

注. 使用時間(分単位)の平均値(*M*)及び中央値(*Mdn*)は、対象の学び方を選択していない者(0分)も分母に含めた値である。

■表12: 「学び方の選択肢（その他を除く）」の観点等別における【選択回数】の記述統計（ $n = 22$ ）

| 段落                      | 学び方の選択肢（観点等別）        | 実践1回目    |            |            | 実践2回目    |            |            |
|-------------------------|----------------------|----------|------------|------------|----------|------------|------------|
|                         |                      | <i>M</i> | <i>Mdn</i> | <i>Max</i> | <i>M</i> | <i>Mdn</i> | <i>Max</i> |
| ①構想                     | 自力 (a)               | 0.95     | 1.00       | 1.00       | 1.00     | 1.00       | 1.00       |
|                         | 内容面 (b), (c),(d),(e) | 0.32     | 0.00       | 1.00       | 0.18     | 0.00       | 2.00       |
|                         | 言語面 (f),(g),(h),(i)  | 0.18     | 0.00       | 1.00       | 0.18     | 0.00       | 1.00       |
|                         | 上記9項目 (a)～(i) 合計     | 1.45     | 1.00       | 3.00       | 1.36     | 1.00       | 4.00       |
| ②執筆                     | 自力 (a)               | 0.86     | 1.00       | 1.00       | 0.64     | 1.00       | 1.00       |
|                         | 内容面 (b), (c),(d),(e) | 0.45     | 0.00       | 2.00       | 0.27     | 0.00       | 1.00       |
|                         | 言語面 (f),(g),(h),(i)  | 0.82     | 1.00       | 2.00       | 0.95     | 1.00       | 2.00       |
|                         | 上記9項目 (a)～(i) 合計     | 2.14     | 2.00       | 4.00       | 1.86     | 2.00       | 4.00       |
| ③修正                     | 自力 (a)               | 0.59     | 1.00       | 1.00       | 0.50     | 1.00       | 1.00       |
|                         | 内容面 (b), (c),(d),(e) | 0.41     | 0.00       | 3.00       | 0.41     | 0.00       | 2.00       |
|                         | 言語面 (f),(g),(h),(i)  | 1.14     | 1.00       | 3.00       | 1.09     | 1.00       | 2.00       |
|                         | 上記9項目 (a)～(i) 合計     | 2.14     | 2.00       | 5.00       | 2.00     | 2.00       | 5.00       |
| 3段階における上記9項目 (a)～(i) 総計 |                      | 5.73     | 5.50       | 11.00      | 5.23     | 5.00       | 9.00       |

注. 個数の平均値(*M*)及び中央値(*Mdn*)は、対象の学び方を選択していない者(0分)も分母に含めた値である。

#### 4.4 学習者間での学び方の選択と使用の違い

最後に本節では、第1節にて明らかになった「複線型」授業において、ライティング得点が〈上昇〉した群、〈不変〉であった群、〈下降〉した群それぞれの学習者個人に焦点を当て、どのような学び方を選択した結果として、そうしたライティングの質の変容に至ったのかを探究する。

まず量的なデータの対応関係から学習者間の共通性を探索的に発見するために、学習者個人(A～X)を、「ライティング得点」が〈上昇〉した群、〈不変〉であった群、〈下降〉した群の3つのグループに分け、学習者個人の事前ライティング得点と事前ライティング方略の対応を表13に、実践2回目での「学び方の選択肢」の観点等別における学び方の【使用時間】の対応を表14に、同じく学び方の【選択回数】の対応を表15にそれぞれ示す。

これら表13から表15を概観したところ、各群の学習者間において確たる共通性は見られなかった。このことから、共通した特定の要因によりライティングの質の変容が引き起こされているのではなく、学習者の個人差要因により多様な方略と学び方が展開されており、それらが複雑に関係し合った結果として、ライティングの質に変容が生じていることが考えられる。

■表13: 学習者個人の事前ライティング得点と事前ライティング方略の対応表

| 群  | 学習者 | 事前ライティング得点 |    |    | 事前ライティング方略 |      |       |        |
|----|-----|------------|----|----|------------|------|-------|--------|
|    |     | 総得点        | 内容 | 言語 | 認知         | メタ認知 | 社会的行動 | 動機づけ調整 |
| 上昇 | A   | 6          | 3  | 3  | 6.42       | 6.17 | 4.83  | 6.00   |
|    | B   | 4          | 2  | 2  | 3.92       | 4.67 | 6.00  | 4.22   |
|    | C   | 3          | 1  | 2  | 3.75       | 5.17 | 2.17  | 4.56   |
|    | D   | 2          | 1  | 1  | 2.25       | 3.50 | 4.17  | 3.11   |
|    | E   | 6          | 3  | 3  | 6.25       | 6.83 | 6.33  | 6.67   |
|    | F   | 3          | 1  | 2  | 5.58       | 5.00 | 5.67  | 5.56   |
|    | G   | 2          | 1  | 1  | 3.50       | 4.67 | 5.00  | 3.89   |
|    | H   | 2          | 1  | 1  | 4.92       | 5.33 | 4.83  | 4.78   |
|    | I   | 3          | 1  | 2  | 3.17       | 3.83 | 6.00  | 3.33   |
|    | J   | 2          | 1  | 1  | 2.75       | 4.17 | 5.17  | 1.67   |
|    | K   | 5          | 3  | 2  | 4.83       | 5.50 | 6.67  | 5.00   |
|    | L   | 9          | 5  | 4  | 4.25       | 4.50 | 4.00  | 4.44   |
|    | M   | 2          | 1  | 1  | 4.83       | 6.00 | 7.00  | 6.44   |
|    | N   | 2          | 1  | 1  | 4.25       | 4.00 | 4.17  | 3.22   |
| 不変 | O   | 4          | 2  | 2  | 2.83       | 4.67 | 5.33  | 3.67   |
|    | P   | 2          | 1  | 1  | 3.75       | 3.83 | 4.83  | 3.22   |
|    | Q   | 4          | 2  | 2  | 5.00       | 6.33 | 6.33  | 5.44   |
|    | R   | 2          | 1  | 1  | 5.08       | 5.83 | 6.00  | 4.44   |
|    | S   | 2          | 1  | 1  | 4.92       | 3.83 | 5.00  | 3.56   |
| 下降 | T   | 6          | 3  | 3  | 4.83       | 3.67 | 4.67  | 4.89   |
|    | U   | 5          | 3  | 2  | 4.92       | 5.50 | 5.50  | 4.33   |
|    | V   | 7          | 4  | 3  | 5.17       | 4.83 | 5.00  | 4.22   |
|    | W   | 6          | 3  | 3  | 3.42       | 5.17 | 4.83  | 4.11   |
|    | X   | 7          | 3  | 4  | 3.42       | 1.00 | 2.17  | 3.56   |

注. ライティング得点は、内容面(0点～5点)、言語面(0点～5点)の2観点、計10点満点である。

■表14: 実践2回目における学習者個人での「学び方の選択肢」の観点等別【使用時間】の対応表

| 群  | 学習者 | ①構想 |    |    |    | ②執筆 |    |    |    | ③修正 |    |    |    |
|----|-----|-----|----|----|----|-----|----|----|----|-----|----|----|----|
|    |     | 自力  | 内容 | 言語 | 計  | 自力  | 内容 | 言語 | 計  | 自力  | 内容 | 言語 | 計  |
| 上昇 | A   | 10  | 0  | 0  | 10 | 2   | 4  | 6  | 12 | 7   | 5  | 10 | 22 |
|    | B   | 5   | 0  | 3  | 8  | 4   | 10 | 10 | 24 | 10  | 0  | 7  | 17 |
|    | C   | 8   | 0  | 0  | 8  | 0   | 0  | 8  | 8  | 0   | 10 | 0  | 10 |
|    | D   | 10  | 0  | 0  | 10 | 0   | 0  | 14 | 14 | 5   | 0  | 10 | 15 |
|    | E   | 11  | 0  | 0  | 11 | 7   | 3  | 0  | 10 | 10  | 5  | 0  | 15 |
|    | F   | 8   | 0  | 0  | 8  | 10  | 12 | 0  | 22 | 0   | 20 | 0  | 20 |
|    | G   | 16  | 0  | 0  | 16 | 0   | 0  | 15 | 15 | 0   | 0  | 20 | 20 |
|    | H   | 9   | 0  | 0  | 9  | 15  | 0  | 5  | 20 | 8   | 0  | 23 | 31 |
|    | I   | 7   | 0  | 0  | 7  | 6   | 0  | 7  | 13 | 5   | 0  | 10 | 15 |
|    | J   | 16  | 0  | 0  | 16 | 0   | 0  | 14 | 14 | 0   | 0  | 17 | 17 |
|    | K   | 10  | 0  | 0  | 10 | 5   | 0  | 5  | 10 | 0   | 0  | 22 | 22 |
|    | L   | 5   | 0  | 0  | 5  | 10  | 0  | 10 | 20 | 3   | 7  | 15 | 25 |
|    | M   | MV  | MV | MV | MV | MV  | MV | MV | MV | MV  | MV | MV | MV |
|    | N   | 5   | 5  | 0  | 10 | 0   | 0  | 5  | 5  | 0   | 0  | 5  | 5  |
| 不変 | O   | 5   | 5  | 5  | 15 | 0   | 0  | 10 | 10 | 2   | 5  | 23 | 30 |
|    | P   | 7   | 0  | 0  | 7  | 8   | 0  | 7  | 15 | 0   | 0  | 25 | 25 |
|    | Q   | 5   | 5  | 0  | 10 | 7   | 0  | 5  | 12 | 2   | 0  | 10 | 12 |
|    | R   | 8   | 0  | 0  | 8  | 5   | 5  | 0  | 10 | 5   | 5  | 15 | 25 |
|    | S   | 10  | 0  | 0  | 10 | 0   | 0  | 10 | 10 | 0   | 0  | 20 | 20 |
| 下降 | T   | 9   | 0  | 2  | 11 | 10  | 0  | 0  | 10 | 0   | 0  | 18 | 18 |
|    | U   | 10  | 0  | 5  | 15 | 5   | 0  | 10 | 15 | 5   | 0  | 10 | 15 |
|    | V   | 10  | 0  | 0  | 10 | 0   | 13 | 0  | 13 | 0   | 16 | 0  | 16 |
|    | W   | 12  | 0  | 0  | 12 | 8   | 0  | 8  | 16 | 0   | 0  | 12 | 12 |
|    | X   | MV  | MV | MV | MV | MV  | MV | MV | MV | MV  | MV | MV | MV |

注. MVは、記録データが欠損していること(Missing Value)を示す。

■表15: 実践2回目における学習者個人での「学び方の選択肢」の観点等別【選択回数】の対応表

| 群  | 学習者 | ①構想 |    |    |    | ②執筆 |    |    |    | ③修正 |    |    |    |
|----|-----|-----|----|----|----|-----|----|----|----|-----|----|----|----|
|    |     | 自力  | 内容 | 言語 | 計  | 自力  | 内容 | 言語 | 計  | 自力  | 内容 | 言語 | 計  |
| 上昇 | A   | 1   | 0  | 0  | 1  | 1   | 1  | 1  | 3  | 1   | 1  | 1  | 3  |
|    | B   | 1   | 0  | 1  | 2  | 1   | 1  | 2  | 4  | 1   | 0  | 2  | 3  |
|    | C   | 1   | 0  | 0  | 1  | 0   | 0  | 1  | 1  | 0   | 1  | 0  | 1  |
|    | D   | 1   | 0  | 0  | 1  | 0   | 0  | 2  | 2  | 1   | 0  | 1  | 2  |
|    | E   | 1   | 0  | 0  | 1  | 1   | 1  | 0  | 2  | 1   | 1  | 0  | 2  |
|    | F   | 1   | 0  | 0  | 1  | 1   | 1  | 0  | 2  | 0   | 1  | 0  | 1  |
|    | G   | 1   | 0  | 0  | 1  | 0   | 0  | 1  | 1  | 0   | 0  | 1  | 1  |
|    | H   | 1   | 0  | 0  | 1  | 1   | 0  | 1  | 2  | 1   | 0  | 1  | 2  |
|    | I   | 1   | 0  | 0  | 1  | 1   | 0  | 1  | 2  | 1   | 0  | 1  | 2  |
|    | J   | 1   | 0  | 0  | 1  | 0   | 0  | 1  | 1  | 0   | 0  | 2  | 2  |
|    | K   | 1   | 0  | 0  | 1  | 1   | 0  | 1  | 2  | 0   | 0  | 2  | 2  |
|    | L   | 1   | 0  | 0  | 1  | 1   | 0  | 2  | 3  | 1   | 2  | 2  | 5  |
|    | M   | MV  | MV | MV | MV | MV  | MV | MV | MV | MV  | MV | MV | MV |
|    | N   | 1   | 1  | 0  | 2  | 0   | 0  | 1  | 1  | 0   | 0  | 1  | 1  |
| 不変 | O   | 1   | 2  | 1  | 4  | 0   | 0  | 1  | 1  | 1   | 1  | 1  | 3  |
|    | P   | 1   | 0  | 0  | 1  | 1   | 0  | 2  | 3  | 0   | 0  | 2  | 2  |
|    | Q   | 1   | 1  | 0  | 2  | 1   | 0  | 1  | 2  | 1   | 0  | 1  | 2  |
|    | R   | 1   | 0  | 0  | 1  | 1   | 1  | 0  | 2  | 1   | 1  | 1  | 3  |
|    | S   | 1   | 0  | 0  | 1  | 0   | 0  | 1  | 1  | 0   | 0  | 1  | 1  |
| 下降 | T   | 1   | 0  | 1  | 2  | 1   | 0  | 0  | 1  | 0   | 0  | 1  | 1  |
|    | U   | 1   | 0  | 1  | 2  | 1   | 0  | 1  | 2  | 1   | 0  | 1  | 2  |
|    | V   | 1   | 0  | 0  | 1  | 0   | 1  | 0  | 1  | 0   | 1  | 0  | 1  |
|    | W   | 1   | 0  | 0  | 1  | 1   | 0  | 1  | 2  | 0   | 0  | 2  | 2  |
|    | X   | MV  | MV | MV | MV | MV  | MV | MV | MV | MV  | MV | MV | MV |

注. MVは、記録データが欠損していること(Missing Value)を示す。

次に、得点が〈上昇〉した群、〈不変〉であった群、〈下降〉した群より、それぞれ特徴的なデータの値を表した学習者個人を抽出し、その学び方の選択と使用を一事例として取り上げ、質的な分析を試みる。

まず、〈上昇〉群より一事例として取り上げるのは、ライティング得点の差分において最も高い値を示していた【学習者A】である。学習者Aの得点は、事前テストで内容面が3点、言語面が3点、総得点が6点であったが、事後テストでは内容面が5点、言語面が5点、総得点が10点満点となり、差分は言語面で2点、内容面で2点、総得点で4点の伸びを示した唯一の学習者であった。学習者Aは、実践1回目では、①自力のみ(10分)で構想し、②言語面について辞書機能・機械翻訳で確認(10分)しながら部分的に自力(8分)で執筆し、③修正では自力(5分)も働かせつつ、言語面について辞書機能・機械翻訳で確認(7分)、生成AIに相談(9分)する学び方を選択して使用していた(合計49分/6個)。実践2回目では、1回目と同じく①自力のみ(10分)で構想するものの、②執筆では内容面について生成AIに相談(4分)するとともに、言語面についても生成AIに相談(6分)し、自力(2分)での執筆はわずかに留め、③修正では自力(7分)も働かせつつ、内容面について教科書・授業資料を参照(5分)したり、言語面について改めて生成AIに相談(10分)したりする学び方を選択して使用していた(合計44分/7個)。実践2回目の学び方の理由について尋ねた追加質問紙では、①構想では「どのような内容を書くかを考えたときに、自分の言葉で伝えるとなると話す順序や構成も変わらと思ったので、自分の力のみで行いました」という理由を、②執筆では「構想したものを自分の力だけで書くのは難しく、どのような表現を用いれば良いか分からなかったので、生成AIを利用しました」という理由を、③修正では「自分の力で調整しつつ、伝えたいことがちゃんと内容に入っているか不安だったため、より説得力のある文にするために教科書や生成AIを参考にしました」という理由をそれぞれ記述していた。

続いて〈不変〉群より一事例として取り上げるのは、各プロセスで使用した「学び方の選択肢」の総個数が多く、費やした時間も長かった【学習者O】である。学習者Oの得点は、事前・事後テストともに内容面が2点、言語面が2点、総得点が4点であった。学習者Oは、実践1回目において、①構想では自力(5分)かつ、内容面について教科書・授業資料を参照(2分)したり、言語面について辞書機能・機械翻訳で確認(8分)したりして構想し、②執筆でも改めて言語面について辞書機能・機械翻訳で確認(6分)しながら部分的に自力(4分)で執筆し、③修正でも自力(12分)を働かせつつ、再び言語面について辞書機能・機械翻訳で確認(18分)する学び方を選択して使用していた(合計55分/7個)。実践2回目においては、①構想では自力(5分)かつ、内容面について教科書・授業資料を参照(2分)、インターネット検索で確認(3分)したり、言語面について辞書機能・機械翻訳で確認(5分)したりして構想し、②執筆では自力を用いずに改めて言語面について辞書機能・機械翻訳で確認(10分)して執筆し、③修正ではわずかに自力(2分)を働かせつつ、改めてインターネット検索で確認(5分)したり、再び言語面について辞書機能・機械翻訳で確認(23分)する学び方を選択して使用していた(合計55分/8個)。実践2回目の学び方の理由について尋ねた追加質問紙では、①構想では「力を伸ばすためにa.(自力)とb.(教科書・授業資料)の時間を多くとった。分からない部分をc.(インターネット検索)で補った」という理由を、②執筆では「内容を考えることはできたが、それを英文にする能力がなかったので、翻訳を利用して書いた」という理由を、③修正では「違和感があったところの文法をインターネットで検索した。文法は翻訳アプリでは分からないため」という理由をそれぞれ記述していた。

最後に〈下降〉群より一事例として取り上げるのは、認知方略の一種であるテキスト処理方略の使用が事前と事後で最も減少していた【学習者V】である。学習者Vの得点は、事前テストで内容面が4点、言語面が3点、総得点が7点であったが、事後テストでは内容面が3点、言語面が2点、総得点が5点となり、事前・事後テストの差分は言語面で-1点、内容面で-1点、総得点で-2点を示した。また、学習者Vのテキスト処理方略の尺度得点の平均値は、事前質問紙で5.00であったが、事後質問紙では3.33となり、差分は-1.67を示した。学習者Vは、実践1回において、①自力のみ(6分)で構想し、②内容面について教科書・授業資料を参照(13分)しながら部分的に自力(3分)で執筆し、③修正では自力は働かせずに、内容面について生成AIに相談(29分)する学び方を選択して使用していた(51分/4個)。実践2回目においては、①自力のみ(10

分)で構想し, ②執筆では自力を用いずに内容面について生成AIに相談(13分)して執筆し, ③修正でも自力は働かせずに, 再び内容面について生成AIに相談(16分)する学び方を選択して使用していた(39分/3個)。実践2回目の学び方の理由について尋ねた追加質問紙では, ①構想では「ネットとかで検索して変な文になってしまうことがあるから」という理由を, ②執筆では「自分で書くより調べたほうが確実だから」という理由を, ③修正では「自分で直すより調べたほうが確実だから」という理由をそれぞれ記述していた。

本節で取り上げた学習者個人における学び方の選択と使用の一事例を整理すると, 次の通りである。

〈上昇〉群より抽出した【学習者A】は, 事前テストの段階ですでに内容面, 言語面ともにライティング得点が到達度の水準に達していたが, 事後テストでは最上位の得点まで上昇していた。そうした得点の上昇を見せた学習者Aが, 実践2回目で選択して使用した学び方は, 構想は自力のみで行い, 執筆から内容面と言語面について生成AIに相談し, 自力での執筆はわずかに留め, 修正では自力を働かせつつ内容面について教科書・授業資料を参照したり, 言語面について改めて生成AIに相談したりする学び方であった。学習者Aが選好した学び方の選択と使用の特徴を端的に表現すると, 構想は自力のみで行い, 執筆, 修正では, それぞれ言語面と内容面の観点から複数の学び方を選択して使用している点が特徴として抽出された。

〈不変〉群より抽出した【学習者O】は, 事前テストでは得点が到達度の水準にはやや達しておらず, 事後テストでも得点は内容面, 言語面とも変容が見られなかった。そうした得点の不変を示した学習者Oが, 実践2回目で選択して使用した学び方は, 構想は自力に加え内容面について教科書・授業資料を参照, インターネット検索で確認したり, 言語面について辞書機能・機械翻訳で確認したりして構想し, 執筆では自力を用いずに改めて言語面について辞書機能・機械翻訳で確認して執筆し, 修正ではわずかに自力を働かせつつ, インターネット検索で文法事項を確認したり, 再び言語面について辞書機能・機械翻訳で確認する学び方であった。学習者Oが選好した学び方の選択と使用の特徴を端的に表現すると, 構想からすでに複数の学び方を選択して使用しており, 執筆, 修正までにかけて繰り返し言語面の観点から辞書機能・機械翻訳で確認しており, 言語面の観点から単一の学び方を繰り返し選択して使用している点が特徴として抽出された。

〈下降〉群より抽出した【学習者V】は, 事前テストでは言語面が到達度の水準, 内容面が到達度をやや上回っていたが, 事後テストでは言語面が到達度をやや下回り, 内容面は到達度から下降していた。また, ライティング方略のうち認知方略の一種であるテキスト処理方略に大きな下がり幅が見られていた。そうした得点と方略の下降を示した学習者Vが, 実践2回目で選択して使用した学び方は, 構想は自力のみで行い, 執筆では自力を用いずに内容面について生成AIに相談して行い, 修正でも自力は働かせずに再び内容面について生成AIに相談する学び方を選択して使用していた。学習者Oが選好した学び方の選択と使用の特徴を端的に表現すると, 構想のみ自力で行い, 執筆, 修正では, 自力は用いずそれぞれ内容面の観点から生成AIと相談に終始しており, 内容面の観点から単一の学び方を繰り返し選択して使用している点が特徴として抽出された。

## 5 総括と展望

本実践の結果と発見は, (RQ1)プロセス・ライティング活動における「複線型」という授業形態は, 「単線型」や「準複線型」に比べて, 学習者集団のライティングの質やライティング方略の変容に有効なのか, (RQ2)「複線型」授業では, 学習者集団の学び方の選択と使用にどのような傾向が見られ, また学習者個人の学び方の選択と使用には, 学習者間でどのような違いが見られるのかという2つの研究課題に答える形で, 主に次の4点にまとめられる:

(1) プロセス・ライティング活動による「複線型」授業は, 「単線型」授業と比べると学習者集団におけるライティングの質の向上に一定の効果がある。

(2)「複線型」授業では、学習者集団におけるライティング方略の使用が一部を除いて減少する傾向にある。

(3)本実践に参加した学習者に選好された学び方の選択と使用は、構想は主に自力で行い、執筆は主に言語面に関する学び方を選択しながら自力で行い、修正でも主に言語面に関する学び方に多くの時間を割く傾向にある。

(4)ライティングの質が向上していない学習者の一部には、単一の観点から単一の学び方を繰り返し選択して使用している特徴が見られる。

以下それぞれについて考察する。第一に、プロセス・ライティング活動による「複線型」授業は、「単線型」授業と比べると学習者集団におけるライティングの質の向上に一定の効果があることについて、「複線型」では学習者が学び方の選択と使用を自己調整する機会を設定したことが、学習者の「個別最適な学び」の実現に寄与し、ライティングの質の向上につながったことが考えられる。この点は「準複線型」の授業形態においても「複線型」に準じた結果が得られたことから、学習者が学び方の選択と使用を自己調整する機会を少なくとも「修正」する段階において設定することは有効であると考えられる。一方で、「構想」および「執筆」する段階においても学習者が学び方の選択と使用を自己調整する機会を設定するか否かは、学習者集団の到達度によって授業形態の適性が異なってくる場合や、授業の回数を重ねる中で段階的に「準複線型」から「複線型」に移行していく方が効果を発揮する可能性なども考えられる。今後対象クラスに適した複線型の授業デザインを探究していきたい。

第二に、「複線型」授業では、学習者集団におけるライティング方略の使用が一部を除いて減少する傾向にあるが、「複線型」授業では、学習者が複数の学び方から選択して使用していたことから、それら複数の学び方に付随する形で使用するライティング方略も分散され、特定のライティング方略の使用の増大には寄与しなかった可能性が考えられる。また、Mizumoto(2018)は学習方略の選択と使用は自己調整能力とメタ認知的知識によって引き起こされると整理しているが、本実践では学習者のメタ認知的知識を担保する手立てが不十分であった。今後はプロセス・ライティング活動の前後の時間や、前後の授業において学習者が計画、実行、振り返りを行う機会を設定し、学習者のメタ認知的知識の蓄積につなげたい。加えて、今回の授業形態でも使用の減少が確認されたピア学習方略やフィードバック操作方略といった社会的行動方略は、「協働的な学び」を実現する上で非常に重要な方略であると考えられる。「個別最適な学び」と「協働的な学び」の一体的な充実を図るために、社会的行動方略の使用を促すような指導者からの働きかけを展開していくことが、今後の改善の方向性として期待される。

第三に、本実践に参加した学習者に選好された学び方の選択と使用は、構想は主に自力で行い、執筆は主に言語面に関する学び方を選択しながら自力で行い、修正でも主に言語面に関する学び方に多くの時間を割く傾向にあることについて、これらのプロセスは対象校の学生たちにとって最も慣れ親しんできた学習スタイルであることが考えられる。こうした学習スタイルは、授業外の機会においても個別に取り組むことができる学び方である点は意義深い。一方で、前述した「協働的な学び」の視点から授業内であるからこそ発揮できる学び方、例えば、他の学生と内容面について相談するような学び方も、学習者が進んで選択していけるような授業デザインが望まれる。

第四に、ライティングの質が向上していない学習者の一部は、単一の観点から単一の学び方を繰り返し選択して使用している特徴が見られる。この点については、あくまで一事例を通した質的な分析結果であることに留意したいが、鈴木(2024)が指摘する学習スタイルの拡張の意義に重なる特徴であると考えられる。特に本実践からは、教科書や生成AIといった複数のメディアやツールを単に用いるのではなく、内容面や言語面といった複数の観点から用いることの重要性が示唆された。本研究では、学習者集団としては実践1回目から2回目にかけて学び方の選択個数が減少する傾向が見られたが、今後さらに3回目、4回目と実践を重ねるにつれて、全ての学習者において学習スタイルが拡張していくような授業デザインを検討していきたい。

## 謝辞

本研究を実施する機会を与えてくださいました公益財団法人 日本英語検定協会および関係者の皆さま、選考委員の先生方に深く御礼申し上げます。中でも、研究助言者の西垣知佳子先生には、授業実践の具体的な構想から論文の執筆にいたるまで、非常に貴重で有益なご助言、ご指導を賜りましたこと、心より深く感謝申し上げます。最後になりますが、授業実践に参加し協力いただいた対象校の学生をはじめ、本研究の実施にご協力いただいたすべての皆さまに心より御礼申し上げます。なお、本報告書に関して、開示すべき利益相反関連事項はありません。

## 引用文献

- 濱田活仁・新美德康・福光将仁 (2025). 「生成 AI による産出モデル文を活用した英作文修正活動における学生の修正行動」『中部地区英語教育学会紀要』54, 299-306. [https://doi.org/10.20713/celes.54.0\\_299](https://doi.org/10.20713/celes.54.0_299)
- 岩本祐樹 (2021). 「ループリックの事前提示がライティングパフォーマンスに与える影響」*EIKEN BULLETIN*, 32, 85-107. [https://www.eiken.or.jp/center\\_for\\_research/pdf/bulletin/vol32/vol\\_32\\_p85-p107.pdf](https://www.eiken.or.jp/center_for_research/pdf/bulletin/vol32/vol_32_p85-p107.pdf)
- Li, S., Hiver, P., & Papi, M (Eds.). (2022). *The Routledge handbook of second language acquisition and individual differences*. Routledge.
- Mizumoto, A. (2018). On questionnaire use in language learning strategies research. *The Journal of AsiaTEFL*, 15(1), 184-192. <https://doi.org/10.18823/asiatefl.2018.15.1.12.184>
- 水本篤・竹内理. (2008). 「研究論文における効果量の報告のために：基礎的概念と注意点」*英語教育研究*, 31, 57-66. [https://www.mizumot.com/files/EffectSize\\_KELES31.pdf](https://www.mizumot.com/files/EffectSize_KELES31.pdf)
- 文部科学省 (2021). 「学習指導要領の趣旨の実現に向けた個別最適な学びと協働的な学びの一体的な充実に関する参考資料」
- 文部科学省・国立教育政策研究所 (2023). 「令和5年度 全国学力・学習状況調査 報告書 中学校 英語」[https://www.nier.go.jp/23chousakekkahoukoku/report/middle\\_eng.html](https://www.nier.go.jp/23chousakekkahoukoku/report/middle_eng.html)
- 鈴木祐一 (2024). 『あたらしい第二言語習得論：英語指導の思い込みを変える』研究社.
- 高橋純 (2022). 「1人1台端末を活用した高次な資質・能力の育成のための授業に関する検討」『日本教育工学研究報告集』4, 82-89. [https://doi.org/10.15077/jsetstudy.2022.4\\_82](https://doi.org/10.15077/jsetstudy.2022.4_82)
- 竹内理 (2014). 「自律的・継続的学習に関する学習者要因の研究動向と今後の展望」全国英語教育学会第40回研究大会記念特別誌編集委員会(編)『英語教育学の今—理論と実践の統合—』234-238.
- Teng, L. S., & Zhang, L. J. (2016). A questionnaire-based validation of multidimensional models of self-regulated learning strategies. *The Modern Language Journal*, 100(3), 674-701. <https://doi.org/10.1111/modl.12339>
- 津久井貴之 (2024). 「個別最適な学びと協働的な学びの実現～個に応じた指導と英語教師の役割に焦点を当てて～」*TEN*, 54, 1-2. [https://tb.sanseido-publ.co.jp/wp-sanseido/wp-content/uploads/2024/09/TEN54\\_web.pdf](https://tb.sanseido-publ.co.jp/wp-sanseido/wp-content/uploads/2024/09/TEN54_web.pdf)
- 保田幸子 (2024). 『「書く力」の発達—第二言語習得論と第二言語ライティング論の融合に向けて—』くろしお出版.
- 横浜国立大学 (2024). 「令和5年度全国学力・学習状況調査の英語の結果を活用した専門的な分析」[https://www.mext.go.jp/content/20240517-mxt\\_chousa02-000034685\\_011.pdf](https://www.mext.go.jp/content/20240517-mxt_chousa02-000034685_011.pdf)

資料1：「複線型」授業における①構想の段階において学生が使ったワークシート(抜粋).....

(1) 構想する段階

どのような文章にするのか、自由に構想しなさい。その際、学び方・時間を自己調整し、下に記録しなさい。

☆必ず記録☆

「(1) 構想する段階」で選択した学び方・時間を、大まかな分単位で下に記録しなさい。(※小数は切り上げ、未選択は空欄)

(例)「a.自分の力のみ」で約5分、「c.内容・構成についてインターネット検索で確認」しながら約8分、「e.内容・構成について他の学生に相談」しながら約2分

| 選択した学び方                  | 時間(分) | 選択した学び方                 | 時間(分) |
|--------------------------|-------|-------------------------|-------|
| a. 自分の力のみ                | 約 分   | f. 英語表現について教科書・授業資料を参照  | 約 分   |
| b. 内容・構成について教科書・授業資料を参照  | 約 分   | g. 英語表現について辞書機能・機械翻訳で確認 | 約 分   |
| c. 内容・構成についてインターネット検索で確認 | 約 分   | h. 英語表現について生成 AI に相談    | 約 分   |
| d. 内容・構成について生成 AI に相談    | 約 分   | i. 英語表現について他の学生に相談      | 約 分   |
| e. 内容・構成について他の学生に相談      | 約 分   | j. その他( )               | 約 分   |

資料2：「複線型」授業における②執筆, ③修正の段階において学生が使ったワークシート(抜粋).....

(2) 書く段階

あなたの意見を示す文章を、英語で丁寧に書きなさい。その際、学び方・時間を自己調整し、下に記録しなさい。

☆必ず記録☆

「(2) 書く段階」で選択した学び方・時間を、大まかな分単位で下に記録しなさい。(※小数は切り上げ、未選択は空欄)

(例)「a.自分の力のみ」で約5分、「f.英語表現について教科書・授業資料を参照」しながら約8分、「g.英語表現について辞書機能・機械翻訳で確認」しながら約2分

| 選択した学び方                  | 時間(分) | 選択した学び方                 | 時間(分) |
|--------------------------|-------|-------------------------|-------|
| a. 自分の力のみ                | 約 分   | f. 英語表現について教科書・授業資料を参照  | 約 分   |
| b. 内容・構成について教科書・授業資料を参照  | 約 分   | g. 英語表現について辞書機能・機械翻訳で確認 | 約 分   |
| c. 内容・構成についてインターネット検索で確認 | 約 分   | h. 英語表現について生成 AI に相談    | 約 分   |
| d. 内容・構成について生成 AI に相談    | 約 分   | i. 英語表現について他の学生に相談      | 約 分   |
| e. 内容・構成について他の学生に相談      | 約 分   | j. その他( )               | 約 分   |

(3) 修正する段階

英作文を修正しながら書き直しなさい。その際、学び方・時間を自己調整し、下に記録しなさい。

☆必ず記録☆

「(3) 修正する段階」で選択した学び方・時間を、大まかな分単位で下に記録しなさい。(※小数は切り上げ、未選択は空欄)

(例)「a.自分の力のみ」で約5分、「f.英語表現について教科書・授業資料を参照」しながら約8分、「h. 英語表現について生成 AI に相談」しながら約2分

| 選択した学び方                  | 時間(分) | 選択した学び方                 | 時間(分) |
|--------------------------|-------|-------------------------|-------|
| a. 自分の力のみ                | 約 分   | f. 英語表現について教科書・授業資料を参照  | 約 分   |
| b. 内容・構成について教科書・授業資料を参照  | 約 分   | g. 英語表現について辞書機能・機械翻訳で確認 | 約 分   |
| c. 内容・構成についてインターネット検索で確認 | 約 分   | h. 英語表現について生成 AI に相談    | 約 分   |
| d. 内容・構成について生成 AI に相談    | 約 分   | i. 英語表現について他の学生に相談      | 約 分   |
| e. 内容・構成について他の学生に相談      | 約 分   | j. その他( )               | 約 分   |

## 資料3: ライティング評価ルーブリック(岩本, 2021)

|   | 内容面                                                                                                                                                                                  | 言語面                                                                                                                                                                                     |
|---|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 5 | <p><u>理由は説得力がある。説明の流れも非常に良く、わかりやすい。</u></p> <ul style="list-style-type: none"> <li>● 理由は、具体的に非常に良く表現されており、妥当である。また、興味深いものである。</li> <li>● ライティングの構成は、首尾一貫してわかりやすく、整理されている。</li> </ul> | <p><u>文法・語彙、文の作り方が非常に良い。</u></p> <ul style="list-style-type: none"> <li>● 語彙が豊富で、正確かつ効果的に使われている。</li> <li>● 主語・動詞のある文をしっかり作ることができる。</li> <li>● 間違いがほとんどない。</li> </ul>                   |
| 4 | <p><u>理由は明確で十分な説得力がある。説明の流れも良く、わかりやすい。</u></p> <ul style="list-style-type: none"> <li>● 理由は、十分に表現されており、妥当である。</li> <li>● ライティングの構成は、わかりやすく、整理されている。</li> </ul>                       | <p><u>文法・語彙、文の作り方が良い。</u></p> <ul style="list-style-type: none"> <li>● 語彙が豊富で、正確に使われている。</li> <li>● 主語・動詞のある文をしっかり作ることができる。</li> <li>● 細かい間違いは多少あるが、意味に関わる間違いはほとんどない。</li> </ul>        |
| 3 | <p><u>理由はある程度明確で、部分的に説得力がある。説明の流れは、良い箇所がある。</u></p> <ul style="list-style-type: none"> <li>● 理由は、おおむね妥当である。</li> <li>● ライティングの構成は、明確である。</li> </ul>                                  | <p><u>文法・語彙、文の作り方は適切であるが、誤りがいくつかある。</u></p> <ul style="list-style-type: none"> <li>● 最低限の語彙を用い表現できている。</li> <li>● 主語・動詞のある文をおおむね作ることができる。</li> <li>● 意味に関わるものも含め、いくつか間違いがある。</li> </ul> |
| 2 | <p><u>理由はほとんど明確ではなく、説得力もほとんどない。説明の流れもあまり良くない。</u></p> <ul style="list-style-type: none"> <li>● 理由は、妥当である箇所がある。</li> <li>● ライティングの構成は、明確である部分もある。</li> </ul>                          | <p><u>文法・語彙、文の作り方はおおむね正しいが、誤りが多くある。</u></p> <ul style="list-style-type: none"> <li>● 語彙の種類が限られている。</li> <li>● 主語や動詞のない文が多く見られる。</li> <li>● 意味に関わるものも含め、間違いが多くある。</li> </ul>             |
| 1 | <p><u>理由は明確ではなく、説得力がない。説明の流れも悪い。</u></p> <ul style="list-style-type: none"> <li>● 理由は、見当外れである。</li> <li>● ライティングの構成が曖昧で、わかりにくい。</li> </ul>                                           | <p><u>文法・語彙、文の作り方が限られており、誤りが多くある。</u></p> <ul style="list-style-type: none"> <li>● 語彙の種類がかなり限られている。</li> <li>● 主語や動詞がある文がほとんどない。</li> <li>● 意味に関わるものも含め、間違いがかなり多くある。</li> </ul>         |
| 0 | <p>上記に記されている基準に達していない。</p>                                                                                                                                                           | <p>上記に記されている基準に達していない。</p>                                                                                                                                                              |

資料4: ライティング方略を測定する質問項目(Teng & Zhang, 2016をもとに筆者が作成).....

| 上位方略                | 下位方略                   | 項目                                          |
|---------------------|------------------------|---------------------------------------------|
| 認知<br>(9項目)         | テキスト処理<br>(6項目)        | 1. 文章を興味深くするために, 英語表現の工夫をする。                |
|                     |                        | 2. 文法の間違いをチェックする。                           |
|                     |                        | 3. スペルや句読点をチェックする。                          |
|                     |                        | 4. 論理的な一貫性があるか構成をチェックする。                    |
|                     |                        | 5. 文と文がつながっているか, 文章にまとまりがあるかをチェックする。        |
|                     |                        | 6. トピックと内容が明確に表現されているかをチェックする。              |
|                     | 授業記憶<br>(3項目)          | 7. 授業で学習した役立つ単語や表現を, 書いて覚える。                |
|                     |                        | 8. 授業で学習した役立つ単語や表現を, 口に出して覚える。              |
|                     |                        | 9. 授業資料や教科書を, 何度も読み返し記憶に定着させる。              |
| メタ認知<br>(6項目)       | 内容構想<br>(3項目)          | 10. 書く前に, 関連する記事を読んで, 内容を考える。               |
|                     |                        | 11. 書く前に, インターネットに関連する情報を検索し, 内容を考える。       |
|                     |                        | 12. 書く前に, 文章の核となる要素について考え, 内容を考える。          |
|                     | 目標志向<br>(3項目)          | 13. 書く学習をするとき, 目標を達成しているかどうか, 自分の学習状況を確認する。 |
|                     |                        | 14. 英作文の授業で, 知識やスキルの習得度を自己評価する。             |
|                     |                        | 15. 書く学習をするとき, 自分の英作文を向上させるための学習目標を設定する。    |
| 社会的行動<br>(6項目)      | ピア学習<br>(3項目)          | 16. 書くためのアイデアを増やすために, 他の学生に相談する。            |
|                     |                        | 17. 書いた英作文を修正するために, 他の学生に相談する。              |
|                     |                        | 18. 他の学生と協力して英作文を完成させる。                     |
|                     | フィード<br>バック操作<br>(3項目) | 19. 他の学生からのアドバイスに基づいて, 自分の英作文を改善しようとする。     |
|                     |                        | 20. 教師からのアドバイスに基づいて, 自分の英作文を改善しようとする。       |
|                     |                        | 21. AI からのアドバイスに基づいて, 自分の英作文を改善しようとする。      |
| 動機づけ<br>調整<br>(9項目) | 興味喚起<br>(3項目)          | 22. 英作文の学習をもっと楽しくする方法を探す。                   |
|                     |                        | 23. 面白いトピックを選んで英作文の練習をする。                   |
|                     |                        | 24. 英作文の課題を, 自分の興味と結びつけるようにする。              |
|                     | 自己対話<br>(3項目)          | 25. 良い成績を取るために, 書く練習をするよう意識する。              |
|                     |                        | 26. 英語で書く能力を向上させるために, 勉強し続ける必要があると意識する。     |
|                     |                        | 27. 自分がどれだけ学べるか知るために, 英作文の授業で学び続けよう意識する。    |
|                     | 感情抑制<br>(3項目)          | 28. 英作文のテストや課題に取り組むとき, 不安を感じないように意識する。      |
|                     |                        | 29. 英作文をあきらめなくなったとき, 「書き続けよう」と意識する。         |
|                     |                        | 30. 英作文をあきらめなくなったとき, 気分を調節する方法を見つける。        |

第37回 研究助成

**B** 実践部門・報告 V・英語能力向上をめざす教育実践

# 高等学校のライティング指導における 生成 AI の活用とその効果 — 機能的適切さ(Functional Adequacy)に着目して —

研究者: 渡邊 大志 神奈川県／神奈川県立総合教育センター 指導主事

(申請時 神奈川県／神奈川県立藤沢工科高等学校 教諭・横浜国立大学教職大学院 在籍)

《研究助言者: 小泉 利恵》

## 概要

本研究は、高等学校における英語ライティング指導において、生成 AI を活用した学習が学習者のパフォーマンス、とりわけ「機能的適切さ(Functional Adequacy)」に与える効果を検証することを目的とした。神奈川県内の公立高校1年生356名(分析対象214名)を対象に、事前・事後テスト、生成 AI を活用した練習課題、およびアンケート調査を実施した。評価は教師3名と Google Gemini により、機能的適切さの4観点(内容・タスク達成度・理解のしやすさ・一貫性と結束性)に基づき行った。分析の結果、教師評価と AI 評価の間には中程度の相関が認められた一方、AI は一貫して高めに評価する傾向があった。混合計画分散分析では「内容」で時間×練習の交互作用が有意となり、練習あり群のみで事前から事後に有意な向上が見られた。他の観点では時間効果は一部認められたものの、練習効果は限定的であった。アンケート調査では、効率性・即時性・気づきの促進・心理的安心感が肯定的に評価される一方、英語の基礎力不足や AI 依存への懸念も示された。以上より、生成 AI は機能的適切さに基づくライティング指導の補助的ツールとして有用であり、その効果を最大化するには、教師による足場かけと批判的活用を促す指導デザインが必要であることが示唆された。

## 1 はじめに

ChatGPTの2022年11月の公開以降、学校教育における生成 AI の利用可能性について議論が行われ、利活用に関する研究も進展している。文部科学省(2024)は、子どもたちが将来生成 AI のようなツールを使いこなせるようになることの重要性を指摘し、学校教育の中で適切な使用方法を体験的に学ぶことが有効であると述べている。同時に、「人間中心の原則」を強調し、生成 AI は「人間の能力を補助・拡張し、その可能性を広げる有用な道具」(p.7)と位置づけられている。つまり、生成 AI はあくまでも人間の目的を達成するための手段であり、活用自体を目的化すべきではない、ということである。

こうした背景を受け、英語教育の分野でも生成 AI を活用した指導法や学習法が検討されている(門田, 2024など)。特にスピーキングやライティングといった発信技能においては、教師に代わってフィードバックを提供できるという点で、時間的・空間的制約を受けずに自律学習を支援する有効なツールとして注目されている。また、学校現場ではライティング指導における添削の時間確保が困難であることも指摘されており(川村他, 2019)、ライティング指導の十分な実施が妨げられている現状がある。もし生徒自身が生成 AI を活用してライティングのパフォーマンスの改善を行うことができれば、教師の負担軽減と同時に生徒の学習機会の拡充にもつながるだろう。

さらに、現行の高等学校学習指導要領では「思考力・判断力・表現力等」の育成が重視されており、単なる言語形式の正確さだけでなく、目的や場面、状況に応じた適切な言語使用や内容面での表現力を高めることが求められている(文部科学省, 2018)。したがって、学習者のパフォーマンスを内容面から評価できる枠組みの導入も併せて考えていく必要があり、この点において「機能的適切さ (Functional Adequacy)」が有効な構成概念となる可能性がある。

本研究では、高等学校におけるライティング指導に焦点をあて、機能的適切さの枠組みを用いた生成AIの活用可能性を実証的に検討する。

2 先行研究

2.1 生成AIによるライティングの評価とフィードバック

生成AIとライティングの評価についての研究で有名なものは Mizumoto and Eguchi (2023) が挙げられる。ChatGPT の GPT-3.5 による自動ライティング採点 (AWS) の信頼性と精度を、人間と機械学習とで比較した研究である。採点対象のライティングとしては、学習者コーパスである TOEFL 11 から英作文 12,100 本を使用した。これらはすでに評価者によって3段階の評価 (high / medium / low) がつけられており、その評価結果が ChatGPT による採点とどの程度一致するかを検討した。生成AIのプロンプトには、IELTS のライティングのルーブリック (10 段階) に基づいて包括的に 0~9 点をつけるように入力した。結果としては、人間と ChatGPT による採点の一致度は比較的高く、生成AIを併用すれば、教師が内容面など必要な部分に時間を割いて集中できるのではないかと述べられている。

ChatGPT を学習者に使用させる際の手順について、水本 (2025) が先行研究を基に示している (表1)。

表1: 英語ライティング指導においてChatGPTを学習者に使用させる際の手順

| ステップ                   | 具体的活動                                                 |
|------------------------|-------------------------------------------------------|
| (1) 事前準備               | ● 教師に ChatGPT の使い方を説明<br>● 学習者に適切な使用方法を指導             |
| (2) 最初のドラフト作成          | ● 学習者が自力でドラフトを書く                                      |
| (3) ChatGPT によるフィードバック | ● 学習者がドラフトを ChatGPT に入力<br>● 文法・語彙・構成・内容についてアドバイスを受ける |
| (4) ドラフトの改訂            | ● 学習者が ChatGPT の助言を基に修正・改善<br>● 必要に応じて ChatGPT に追加質問  |
| (5) 教師やピアによるレビュー       | ● 改訂版に対してフィードバックを得る                                   |
| (6) 最終版の作成             | ● 受けたフィードバックを踏まえて最終版を完成                               |
| (7) 振り返り               | ● ChatGPT 使用経験や学習効果について学習者が振り返りを行う                    |

注: 水本 (2025) より引用

この手順の中では、学習者は教師の支援を受けながら、ChatGPT を活用して自身のライティングスキルを自律的に向上させる方法を体験的に学ぶことになる。重要なことは、教師の役割であり、教師は生成AIが学習者によって効果的に使用されるように指南するべきである。特に、生成AIからのフィードバックから「気づき (noticing)」が促進されるような手立てや授業デザインをしていくことに、留意すべきであると述べている。

生成AIを活用したライティングの授業の実践例としては山下他(2024)がある。大学生が自身で使用するためのAIを含む英文作成支援ツールTransableを開発し、それを実際に3回の授業で使用した実践報告である。Transableに搭載されているシステムは、以下の4つである。

- Grammarly: 文法や語法の修正提案ができる
- DeepL: 日本語訳の確認ができる(意図した内容になっているか)
- ChatGPT: 様々な表現を調べる(テンプレートでプロンプトが用意されている)
- TOEFL iBT, GTEC, CEFR, IELTS のルーブリック: 選ぶと即時的に採点とコメントを出力する

結果としては、生徒のライティング力の向上も見えて取れ、生徒のTransableに対する使用感は肯定的な意見が多かった。一方で、フィードバックを受けての生徒によるほとんどの修正が、単語レベルでの修正にとどまっている点が懸念点として挙げられていた。生成AIが生徒の自律的な学習ツールとして機能するようになれば、教師がフィードバックする内容をより焦点化したり、個別最適化されたフィードバックを生徒に提供したりすることが可能になるだろうということが示唆された。

学習者による自動採点システムからのフィードバックを活用したライティングの修正に関しては、生徒が修正した量だけでなく、どのような修正を行ったのかという質にも気を配る必要がある。Liu and Yu(2022)は、参加者にWrite & Improve (W & I) with Cambridgeという自動採点システムから得られたフィードバックを基に、自分で書いたものの修正を求めた。W & I with Cambridgeは直接語彙フィードバック(DWF)、間接語彙フィードバック(IWF)、間接文レベルフィードバック(ISF)の3つのフィードバックを行うことができるシステムである。学習者がそれらをどのように活用するかを、アイ・トラッキング、回顧的インタビュー、省察記録を基に分析した。結果は、フィードバックを用いて修正された率は75%と高かったが、表面的なミスの修正が多く、内容面(アイデアや論理構成)に関する修正は23%と少なめであった。直接的なフィードバックは修正につながりやすいが学びにはつながりにくい一方で、間接的なフィードバックは学習者のレベルが一定以上ないと修正につながりにくい可能性が示唆された。これは、W & I with CambridgeがDWFやIWFが中心であることや、ISFが単に「文の質が低い」ことをその部分をハイライトして示すだけである、といった自動採点システムの制約によるものが考えられる。内容面に関する直接的なフィードバックの効果の検討が必要であろう。

さらに近年では、生成AIを含む自動フィードバックや自動評価(AWE/AWF)に関するレビューや研究が多数報告されている。Shi and Aryadoust(2024)は1993年以降のAIベースの自動フィードバック研究を体系的に整理し、AWFは文法や語彙といった表層的誤り訂正には有効である一方、内容や構成など高次スキルへの効果は限定的であることを示した。また、Karatay and Karatay(2024)はAWE研究の統合的レビューを行い、AWEが学習者の自己修正や自律性を促進することを認めつつも、教師フィードバックを完全に代替するものではなく、補完的に活用すべきだと論じている。また、Feng et al.(2025)は2014年から2024年までの112本の研究をレビューし、AIとL2ライティング研究の多くが大学生を対象に短期的効果を検証するもので、理論的基盤が弱い点を指摘している。特に、ChatGPTの無料公開以降の研究では、学習者や教師がAIをどのように受け止めるかや、AIがアイデア生成や協働的学習をどのように支援するか、といった新たな課題が急速に注目されつつあることが報告されている。学習者の認識面については、Allen and Mizumoto(2024)が日本の大学生を対象に、ライティング活動におけるピアレビューとChatGPTからのフィードバックを比較した研究を行い、多くの学習者がAIを「信頼できる権威的なフィードバック源」として評価していたことを明らかにした。同時に、学習者同士の協働的な学びにも一定の価値を感じていることもわかり、AIとピアフィードバックの補完的活用が望ましいと結論づけている。

また、AI評価の信頼性に関しては、Tate et al.(2024)がChatGPTによるエッセイスコアリングを分析し、人間評価者との一致はおおむね中程度であるが、内部的・一貫性(再現性)はむしろAIの方が高い場合もあると報告している。これにより、形成的評価では十分に活用可能である一方で、ハイスタークスの試

験や総括的評価においては、まだ慎重な運用が必要であるとされる。評価の場面でどのように生成AIを導入すべきかどうかに関しては、Voss et al.(2023)が言語テストにおける生成AIの利用についての議論において、実際の言語使用を反映するためにはAI利用を許容すべきだという立場と、構成概念の歪みや公平性・倫理的风险から制限すべきだという立場からの視点や考え方を整理している。そのまとめにおいて、言語テストとAIの開発者が協働し、人間の価値観に合致した適切な支援技術の利用を可能にする「ガードレール」を備えたアセスメントを設計していくことが求められると述べている。

2.2 機能的適切さ

学習者の言語パフォーマンスの内容面についての評価基準として、機能的適切さ (Functional Adequacy) がある。Kuiken and Vedder (2017) は、第二言語ライティング評価における従来の複雑さ (complexity)、正確さ (accuracy)、流暢さ (fluency) といった CAF と呼ばれる指標では、タスク達成度や意味伝達の適切さを十分に捉えられないと指摘し、新たに機能的適切さの評価枠組みを提案した。それまで内容を評価するための評価軸は様々提唱されてきていたが、Kuiken and Vedderは①内容、②タスク達成度、③理解のしやすさ、④一貫性と結束性の4つの観点に整理し、6段階のリッカートスケールで学習者の産出を評価するループリックを作成した(表2)。

■表2: 機能的適切さ評価する4つの観点

|                                                                            |
|----------------------------------------------------------------------------|
| 内容(Content): 文章に含まれる情報単位の数に適切で関連性があるか?                                     |
| 1: アイデアの数がまったく適切でなく不十分であり、アイデア同士に関連性がない。                                   |
| 2: アイデアの数がほとんど適切ではなく、アイデアに一貫性がない。                                          |
| 3: アイデアの数はある程度適切だが、あまり一貫性がない。                                              |
| 4: アイデアの数は適切であり、十分に一貫している。                                                 |
| 5: アイデアの数は非常に適切であり、非常に一貫している。                                              |
| 6: アイデアの数は極めて適切であり、非常に一貫している。                                              |
| タスク達成度(Task requirements): 課題の要件(ジャンル、発話行為、レジスターなど)をどれだけ満たしているか?           |
| 1: 課題の質問および要件のいずれにも答えていない。                                                 |
| 2: 質問および要件のうち、半分未満しか答えていない。                                                |
| 3: 質問および要件のうち、約半分に答えている。                                                   |
| 4: 質問および要件のうち、大半(半分以上)に答えている。                                              |
| 5: ほとんどすべての質問および要件に答えている。                                                  |
| 6: すべての質問および要件に答えている。                                                      |
| 理解のしやすさ(Comprehensibility): 読者が文章の目的や主張を理解するために必要な努力の程度はどれくらいか?            |
| 1: 文章はまったく理解できない。アイデアや目的が明確に示されておらず、読者が文章を理解しようとしても効果がない。                  |
| 2: 文章はほとんど理解できない。目的が明確でなく、読者は筆者のアイデアを理解するのに苦労する。読者はほとんどのアイデアや目的を推測する必要がある。 |
| 3: 文章はある程度理解できる。いくつかの文は一読では理解しにくい。再読によって目的やアイデアが明確になるが、いくつかの疑問は残る。         |
| 4: 文章は理解可能である。ごく少数の文が不明瞭だが、再読によって無理なく理解できる。                                |
| 5: 文章は非常に理解しやすく、読みやすい。理解の問題はまったくない。                                        |
| 6: 文章は極めて理解しやすく、非常に読みやすい。アイデアと目的が明確に示されている。                                |

## 一貫性と結束性 (Coherence and cohesion) : 文章は一貫性と結束性があるか? (結束装置やストラテジーの使用など)

- 1: 文章はまったく一貫性がない。無関係な文の流れや一貫性の破綻が非常に多い。筆者は照応表現をまったく使っておらず、文章はまったく結束していない。接続表現はほとんど使われず、アイデア同士に関連性がない。
- 2: 文章の一貫性は乏しい。筆者はしばしば無関係な文の流れを使っており、一貫性があってもそれは繰り返しによるものである。照応表現はごくわずかにしか使われていない。一貫性の破綻が見られる。文章の結束性も乏しく、アイデアは接続表現でうまくつながっておらず、接続表現自体がほとんど使われていない。
- 3: 文章にはある程度の一貫性があるが、無関係な文の流れや繰り返しが頻繁に見られる。同じ主語が(省略されていても)連続する文で使われることが多い。いくつかの照応表現は使われているが、一貫性の破綻もある。文章にはある程度の結束性があり、一部の接続表現は使われているが、主に接続詞に限られている。
- 4: 文章には一貫性がある。無関係な文の流れは稀にあるが、筆者は時折繰り返しに頼って一貫性を保っている。十分な数の照応表現が使われており、一貫性の破綻は一部あるかもしれないが、文章は結束しており、書き手は接続表現を上手に使っている。使われる接続表現は接続詞に限られないこともある。
- 5: 文章は非常に一貫している。筆者が新しいトピックを導入するとき、それは通常接続詞や接続句を使って行われる。繰り返しは非常に稀である。照応表現は多数使われており、一貫性の破綻はない。文章は非常に結束しており、アイデアは副詞的または動詞的な接続表現によってよく結びつけられている。
- 6: 筆者は、新しいアイデアを接続詞や接続句で統合することで、極めて一貫性を確保している。照応表現は適切に使用されている。無関係な文の流れはほとんどなく、一貫性の破綻はない。文章の構成は非常に結束しており、接続表現(特に接続句、動詞構文、副詞句)を巧みに使用して、アイデア間の関係を記述している。

注. Kuiken and Vedder (2017) より引用, 日本語訳は筆者による

Kuiken and Vedder による、オランダの学習者を対象としたライティング課題の分析では、この新しい尺度が評価者間で高い信頼性を示し、CAF だけでは捉えきれないタスク達成度などを測定できることが確認された。また、CAF とは部分的な関連が見られるものの、必ずしも言語的な正確さが高いことと機能的適切さが一致するわけではなく、誤りがあっても十分に伝達できれば、機能的適切さが高く評価され得ることも示された。さらに、学習者にフィードバックをする際、与えられたトピックに応じたフィードバックを与えることができることも、機能的適切さを評価軸として用いることの利点であると主張している。

機能的適切さが CAF とは異なる指標であるという研究は他にもある。例えば、Révész et al. (2016) は、機能的適切さに類似した指標として、コミュニケーションの適切さ (Communicative adequacy) と CAF の関連を検証し、CAF だけでは学習者の産出の全体的な質を説明するには不十分であると報告している。また、Koizumi and In'nami (2024) も CAF と機能的適切さの関連を調査し、機能的適切さが CAF からある程度予測できることは認めつつも、両者は異なる構成概念であるとし、CAF では説明しきれない学習者のパフォーマンスの側面を機能的適切さが補完すると論じている。De Jong et al. (2007) のタスクの複雑性が流暢性と機能的適切さに与える影響を調査した研究においても、機能的適切さは CAF とは異なる次元の評価であると主張している。以上を踏まえると、これらの先行研究では CAF と機能的適切さが部分的に関連はするものの、別の能力を評価するものであると位置づけている。一方、Lu and Suzukida (2025) では、スピーキングにおいて機能的適切さと複雑性や正確性との関係は見られなかったが、ライティングにおいては複雑性と正確性が高いと機能的適切さが高まったと述べている。また、熟達度の低い学習者は、ライティングにおいて正確性、特に語彙の正確性が機能的適切さと関係があることを報告している。

鈴木 (2023) は、スピーキング能力における機能的適切さについて自動採点の可能性に言及し、研究の必要性について述べているが、ライティング能力における機能的適切さについての自動採点、特に生成 AI を用いている研究はほとんどない。そこで、本研究では、高等学校における英語ライティング指導において、生成 AI から得た機能的適切さに関するフィードバックが学習者のライティングのパフォーマンス

にどのような影響をもたらすかを明らかにし、高等学校の英語ライティング指導における生成AIの効果的な活用方法を検討することを目的とする。以下のリサーチクエスチョンを設定する。

**RQ1** 英語ライティングに対する機能的適切さについての教師の評価とAI評価はどの程度一致するか。

**RQ2** 生成AIによる機能的適切さに関するライティングへのフィードバックは、学習者のパフォーマンスを向上させるか。

**RQ3** 学習者は生成AIによる機能的適切さに関するライティングへのフィードバックについてどのように受け取り、修正に生かしているか。

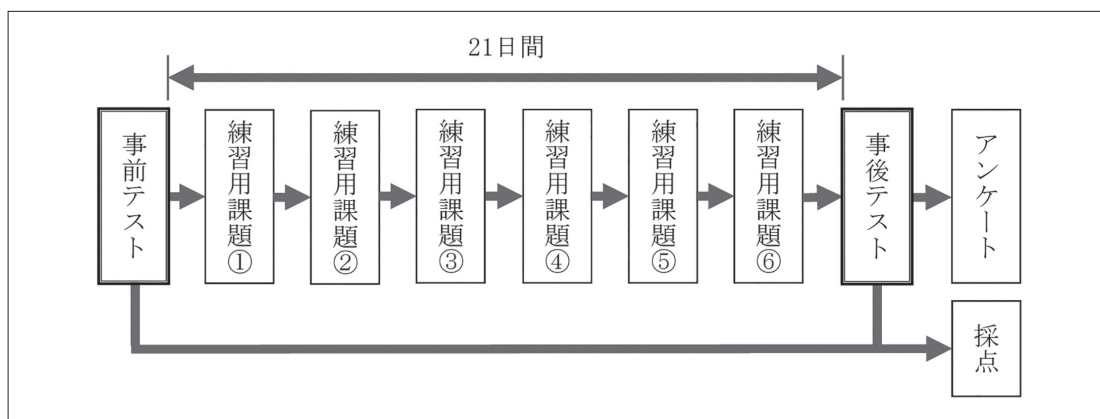
## 3 方法

### 3.1 研究対象者

本研究は、神奈川県内の公立高校1校の第1学年の356名の生徒を対象に実施した。対象者の英語レベルは、CEFRのB1相当である。ほぼ100%の生徒が大学入試の一般受験をするため、英語学習に対するモチベーションは高い。教師は日ごろから教科書で扱われている内容に関連したスピーキングやライティングを行っており、クラスによって多少のばらつきはあるようだが、英語を使って自身の考えを発信する活動に対して概して積極的に取り組んでいる。これまでの英語科の授業において、生成AIを活用したり、活用方法を指導したりしていないとのことだったため、本研究の協力を依頼する対象者として適切であると判断した。研究実施にあたり、まず当該校の校長に研究の目的や内容、データの活用の仕方について口頭と書面で説明をし、承認を得た。加えて、対象生徒に対しても、本研究の目的や方法、得られた情報の取り扱いについて十分な説明を行った。あくまでも自由意思での参加であり、参加しない場合にも不利益はないこと、本研究に係るテストや課題の提出状況などの情報は研究目的以外で使用されることはなく、学校の成績には一切反映されないこと、研究実施期間中であっても参加を辞退できることなどを口頭で伝え、生徒から同意を得た上で実施をした。

### 3.2 データ収集

データ収集は、事前テスト、練習用課題、事後テスト、アンケート、採点という手順で行った(図1)。



■図1: データ収集の手順

### 3.2.1 事前テスト

事前テストは、英語コミュニケーション I の授業内で行われた。テストは、Google フォームで作成し、ライティングのトピックが提示されてから2分間の準備時間を設け、その後10分間でライティングを行う形式で実施した。この形式は、生徒たちが普段から授業内で行っている形式であり、本研究においてもそのまま採用することで、テスト形式に関しての混乱が起きないように配慮した。トピックは“A gift I got from someone”「誰かからもらった贈り物について友人に紹介する」という内容であった。このトピックは、当時行われていた授業の単元におけるゴールタスクのトピック「海外の食べ物をクラスメイトに紹介する」に関連させ、身近な話題に関するものを他人に紹介するというテーマ設定にすることで、授業との関連性を持たせた。なお、事後テストでも同じものを使用したため、事前テストで書いたものに関しては生徒に結果をフィードバックをしていない。

### 3.2.2 練習用課題

事前テストの後、21日間の間に3～4日おきに Google フォームで作成した練習用課題を配信し、授業外での取り組みを依頼した。練習用課題では、まず事前テストと同様に、トピック提示後に2分間の準備と10分間のライティングを行った。その後、書いた英文を、Kuiken and Vedder (2017) の機能的適切さの評価ルーブリックに基づいた評価とフィードバックを要求するプロンプト(筆者作成、図2を参照)と共に無料版の Google Gemini 2.5 Flash (以下、Gemini) に入力し、得られたフィードバックを基にリライトを行った。プロンプト内の指示に関しては、生徒や教師が使用しやすいように、なるべく短く簡潔な指示にとどめた。生徒に課題へ取り組ませる前に、筆者自身が試しに自作の英文を Gemini に入力し、フィードバックを求めてみた。すると、返ってきたコメントは観点別のものではなく全体的に曖昧であり、しかも高い確率で Gemini が英文全体を書き直して提示してしまうことがわかった。そこで、「それぞれの観点に関して改善のためのアドバイスをできるだけ具体的に個別に提示してください」という指示と、「書き直した全体の改善例は出さないでください」という指示を加えた。その結果、Gemini は機能的適切さの観点ごとに、生徒のライティング改善を促すような具体的なフィードバックを出力するようになった。筆者はこのプロンプトを練習用課題で用いるよう、生徒に伝えた。

あなたは英語ライティングの先生です。

生徒が“Something I Use Everyday”というトピックで、他人に紹介する内容のライティングをしました。これから入力されるその英文を、以下のルーブリックにしたがって観点別に評価をし、それぞれの観点に関して改善のためのアドバイスをできるだけ具体的に個別に提示してください。

※書き直した全体の改善例は出さないでください。

【表2の機能的適切さの評価基準をここに入力】

#### ■図2: 練習用課題に使用したプロンプト

生徒は最初にした英文と、Gemini からの評価とフィードバック、リライトしたもの、気づいたことや学んだことや教師に質問したいことを、すべて Google フォームに入力し送信した。Gemini からのフィードバックを用いたこれらの手順については、水本 (2025) の提示した手順 (表1) を参考にしたが、今回は授業外の自律的な学習として練習用課題を作成したため、(5) 教師やピアによるレビューと (6) 最終版の作成の段階は行わなかった。

練習課題のトピックについては、身近な話題に関するトピックを10個、当該生徒を指導している英語教師に提示して意見を仰ぎ、最終的に表3に示した6つに決定した。

■表3: 練習課題のトピック

|   |                                    |
|---|------------------------------------|
| ① | My favorite character              |
| ② | Something I use every day          |
| ③ | A convenient item for school life  |
| ④ | Something that helps me relax      |
| ⑤ | The best place I have ever visited |
| ⑥ | A book or movie that I like        |

なお、当該学年の英語指導を担当している教師に、クラウドサービス上での定期的な課題の配布と、課題実施の生徒への声かけを、期間中毎日してもらった。生徒には、課題について配信された順番に取り組む必要はなく、取り組めるタイミングで自身で選んだ課題に取り組むように伝えた。

### 3.2.3 事後テスト

事後テストは、事前テストと同じ条件・トピックで行った。生徒は事前テストと全く同じものであることは知らされていなかった。採点は事前テストと同じ基準・方法で行った。事前テスト同様、事後テストにおいて書いたものについても、Geminiからのフィードバックは受けていない。

### 3.2.4 事後アンケート

事後テストの実施後に、対象者がAIと機能的適切さに関することをどのように受け取っているかを調査するため、Googleフォームを用いてアンケートを実施した。アンケートの項目は、①AIを使ったライティングのフィードバックを使ってみてどう感じたか、②4つの観点(内容・理解のしやすさ・タスク達成度・一貫性と結束性)に基づいたフィードバックについてどのように感じたか、③その他感想やコメント、とした。実際にライティングの学習に使用した率直な感想を書いてもらえるよう、各項目の質問は大まかな内容にとどめ、自由記述形式での回答を求めた。授業時数の関係で、アンケートは授業外での実施となった。

### 3.2.5 採点

テストの採点は、著者と現職の高校教師1名、教職大学院在籍で英語科教師志望の大学院2年生1名の計3名と、無料版のGeminiで行った(採点時期は2025年7月下旬)。今回Geminiを採用した理由は、当該高等学校を所管する神奈川県がGoogleを学習のインフラとして採用しており、全ての生徒にGoogleのアカウントが配布されていることから、本研究実施後の使用継続の可能性も考慮したためである。また、無料版を選んだのも、すべての生徒が使用可能なツールを用いた場合の効果を検討するためである。

評価基準にはKuiken and Vedder(2017)の機能的適切さ(4観点×6段階)のスケールをそのまま採用した。Koizumi and In'nami(2022)は、教室内のスピーキングテストを用いて、機能的適切さの評価スケールを検討する研究を行ったが、評価スケールは全体的に機能していた一方で、評価者間の一致や、高いレベルの評価に少し問題があったため、1つの解決策として評価スケールを3段階にすることを提案している。しかし、Mizumoto and Eguchi(2023)はIELTSの10段階のルーブリックを用いてAIによるスコアリングを実施し、同じライティングについて2回評価をした際の一致率は52%、1ポイント差だったものは44%であったという点を踏まえ、生成AIによる評価に多少ブレはあるものの、この程度のばらつきは許容範囲であると判断し、評価のスケールを縮小せずそのまま用いた。

採点前には、事後テストが欠損している6人の生徒の事前テストのライティングを練習用として用いて、採点者3名で採点基準のすり合わせを行った。そこで合意した6つのライティングに対する評価と、機能的適切さのスケールをプロンプトとともに Gemini に入力し、採点するように指示をした(図3)。採点は、事後テストが終わってから行った。採点時には、個人が特定されないように処理をし、事前テストと事後テストの組み合わせもわからないように配慮した。

あなたは英語ライティングの先生です。

生徒が“My favorite character”というトピックで、他人に紹介する内容のライティングをしました。

これから入力されるその英文を、以下のルーブリックにしたがって観点別に評価をしてください。

【表2の機能的適切さの評価基準をここに入力】

お送りしたファイルに採点例がありますので、それを参考に採点をお願いします。

採点結果の数字のみ教えてください。

■図3: 採点に使用したプロンプト

### 3.3 データ分析

RQ1については、教師の評価と生成AIの評価の一致度を検討するために、教師と生成AIによる事前・事後テスト両方の得点を、機能的適切さの4つの観点別にそれぞれクロンバックの $\alpha$ 係数を算出した。その後、教師の評価と生成AIの評価について、スピアマンの相関分析を行い、相関係数を算出した。

RQ2に関しては、生成AIを用いたライティングの自律的学習による、生徒のパフォーマンスについての機能的適切さに対する効果を検討するために、機能的適切さの4つの観点別それぞれについて、事前・事後テストの得点を、練習あり・なしのグループに分けて比較、検討した。前提検討として、Levene 検定を行い、群間等分散性(内容:事前  $p=.11$ , 事後  $p=.23$ ; タスク達成度:事前  $p=.13$ , 事後  $p=.76$ ; 理解のしやすさ: 事前  $p=.07$ , 事後  $p=.95$ ; 一貫性と結束性:事前  $p=.27$ , 事後  $p=.10$ )と、残差の Q-Q プロットをそれぞれ確認し正規性はおおむね良好であったため、2要因の混合計画分散分析を用いて分析を行った。

RQ3に関して、生成AIによる機能的適切さに関するライティングへのフィードバックを生徒がどのように受け取り、修正に活かしているかを見るために、生徒からのコメントとリライトされたライティングを質的に分析した。

なお、統計的な分析についてはJASP (Version 0.95)を使用した。

## 4 結果

授業のスケジュールの関係で、全9クラスのうち1クラスは練習のための期間が十分に取れず、もう1クラスは事後テストを実施する時間が確保することができなかったため、当該のクラスの生徒のデータは分析から除外し、対象者は276名になった。さらに、授業への欠席や遅刻等で事前・事後テストの両方、もしくはいずれかを受験できなかった生徒62名も分析から除外した。最終的に、214名分のデータを分析対象とした。授業外における練習課題に一度でも取り組んだ生徒を練習ありグループ( $n = 62$ )、一度も取り組まなかった生徒を練習なしグループ( $n = 152$ )としてグループ分けをし、結果を比較した。生徒の練習課題に対する取り組み状況は表4に示した。

■表4: 生徒の練習課題への取り組み状況

| 練習回数 | 生徒の人数 | 練習課題 | 取り組んだ生徒の人数 |
|------|-------|------|------------|
| 0    | 152   | ①    | 58         |
| 1    | 37    | ②    | 47         |
| 2    | 16    | ③    | 12         |
| 3    | 5     | ④    | 16         |
| 4    | 4     | ⑤    | 5          |
| 5    | 0     | ⑥    | 0          |

#### 4.1 教師の評価と生成AIの評価の一致度 (RQ1)

まず、教師の3人の評価者の一致度を検討するために、事前・事後テストの結果における機能的適切さのそれぞれの観点について、3人のクロンバックの $\alpha$ 係数をそれぞれ算出した(表5)。

■表5: 評価者3人の記述統計と信頼性係数(クロンバック $\alpha$ )

| 観点      | テスト | N   | 評価者A |      | 評価者B |      | 評価者C |      | 評価者3人 |      |          |
|---------|-----|-----|------|------|------|------|------|------|-------|------|----------|
|         |     |     | M    | SD   | M    | SD   | M    | SD   | M     | SD   | $\alpha$ |
| 内容      | 事前  | 214 | 3.71 | 1.02 | 3.32 | 0.88 | 3.72 | 0.81 | 3.56  | 0.75 | 0.76     |
|         | 事後  | 214 | 4.14 | 0.96 | 3.62 | 0.77 | 3.79 | 0.61 | 3.85  | 0.62 | 0.68     |
| タスク達成度  | 事前  | 214 | 3.42 | 1.05 | 3.57 | 0.86 | 4.01 | 0.79 | 3.67  | 0.72 | 0.71     |
|         | 事後  | 214 | 4.01 | 1.04 | 3.80 | 0.44 | 4.16 | 0.67 | 3.99  | 0.67 | 0.74     |
| 理解のしやすさ | 事前  | 214 | 3.57 | 0.99 | 3.76 | 0.60 | 3.73 | 0.81 | 3.69  | 0.62 | 0.64     |
|         | 事後  | 214 | 4.16 | 1.02 | 3.86 | 0.44 | 3.72 | 0.80 | 3.92  | 0.56 | 0.50     |
| 一貫性と結束性 | 事前  | 214 | 3.15 | 1.07 | 3.13 | 0.79 | 3.43 | 0.76 | 3.23  | 0.70 | 0.71     |
|         | 事後  | 214 | 3.77 | 1.06 | 3.24 | 0.54 | 3.65 | 0.60 | 3.23  | 0.70 | 0.73     |

評価者3名の内的一貫性は、内容(事前 $\alpha = 0.76$ , 事後 $\alpha = 0.68$ )、タスク達成度(事前 $\alpha = 0.71$ , 事後 $\alpha = 0.74$ )、一貫性と結束性(事前 $\alpha = 0.71$ , 事後 $\alpha = 0.73$ )においては良好、理解のしやすさ(事前 $\alpha = 0.64$ , 事後 $\alpha = 0.50$ )は中程度であった。観点によって多少ばらつきがあるものの、総じて一定の信頼性が確認されたため、以降の分析では3名の評価の平均値を教師評価として用いた。

次に、教師とGeminiによる採点の一致度を検討するために、教師の採点結果とGeminiによる採点結果を、機能的適切さの4つの観点別にスピアマンの相関分析を用いて比較をした(表6)。内容については、教師の採点結果( $M = 3.69$ ,  $SD = 0.83$ )とGeminiの採点結果( $M = 4.99$ ,  $SD = 1.00$ )には両者の間に中程度の正の相関が認められた( $r = .58$ ,  $p < .001$ )。タスク達成度についても、教師の採点結果( $M = 3.70$ ,  $SD = 0.83$ )とGeminiの採点結果( $M = 5.14$ ,  $SD = 0.94$ )の間に中程度の正の相関が確認された( $r = .54$ ,  $p < .001$ )。さらに、理解のしやすさでは教師の採点結果( $M = 3.84$ ,  $SD = 0.65$ )とGeminiの採点結果( $M$

= 4.52,  $SD = 0.76$ )の間に中程度の正の相関が認められた( $r = .48, p < .001$ )。一貫性と結束性についても、教師の採点結果( $M = 3.32, SD = 0.79$ )と Gemini の採点結果( $M = 4.94, SD = 1.10$ )の間で中程度の正の相関が確認された( $r = .51, p < .001$ )。このように、すべての観点において、教師と Gemini の採点結果には中程度の相関が見られた。しかしながら、教師の採点結果に比べると、Gemini の採点結果の方が0.68～1.62点高くなっていた。有意差の有無を確認するために、まずシャピロ・ウィルク検定を行い、すべての観点において正規性が棄却( $p < .001$ )されたので、ウィルクソンの符号順位検定を用いて比較した。結果は、4観点すべてで有意差が確認された(内容: $z = 16.88, p < .001, r_{rb} = .98$ ;タスク達成度: $z = 17.34, p < .001, r_{rb} = .99$ ;理解のしやすさ: $z = 14.48, p < .001, r_{rb} = .90$ ;一貫性と結束性: $z = 17.09, p < .001, r_{rb} = .97$ )

■表6: 教師とGeminiによる評価の相関および差の検定

| 観点      | 教師   |      | Gemini |      | スピアマンの相関分析 |           | ウィルクソンの符号順位検定 |           |          |
|---------|------|------|--------|------|------------|-----------|---------------|-----------|----------|
|         | $M$  | $SD$ | $M$    | $SD$ | $r$        | $p$       | $z$           | $p$       | $r_{rb}$ |
| 内容      | 3.69 | 0.83 | 4.99   | 1.00 | .58        | < .001*** | 16.88         | < .001*** | .98      |
| タスク達成度  | 3.70 | 0.83 | 5.14   | 0.94 | .54        | < .001*** | 17.34         | < .001*** | .99      |
| 理解のしやすさ | 3.84 | 0.65 | 4.52   | 0.76 | .48        | < .001*** | 14.48         | < .001*** | .90      |
| 一貫性と結束性 | 3.32 | 0.79 | 4.94   | 1.10 | .51        | < .001*** | 17.09         | < .001*** | .97      |

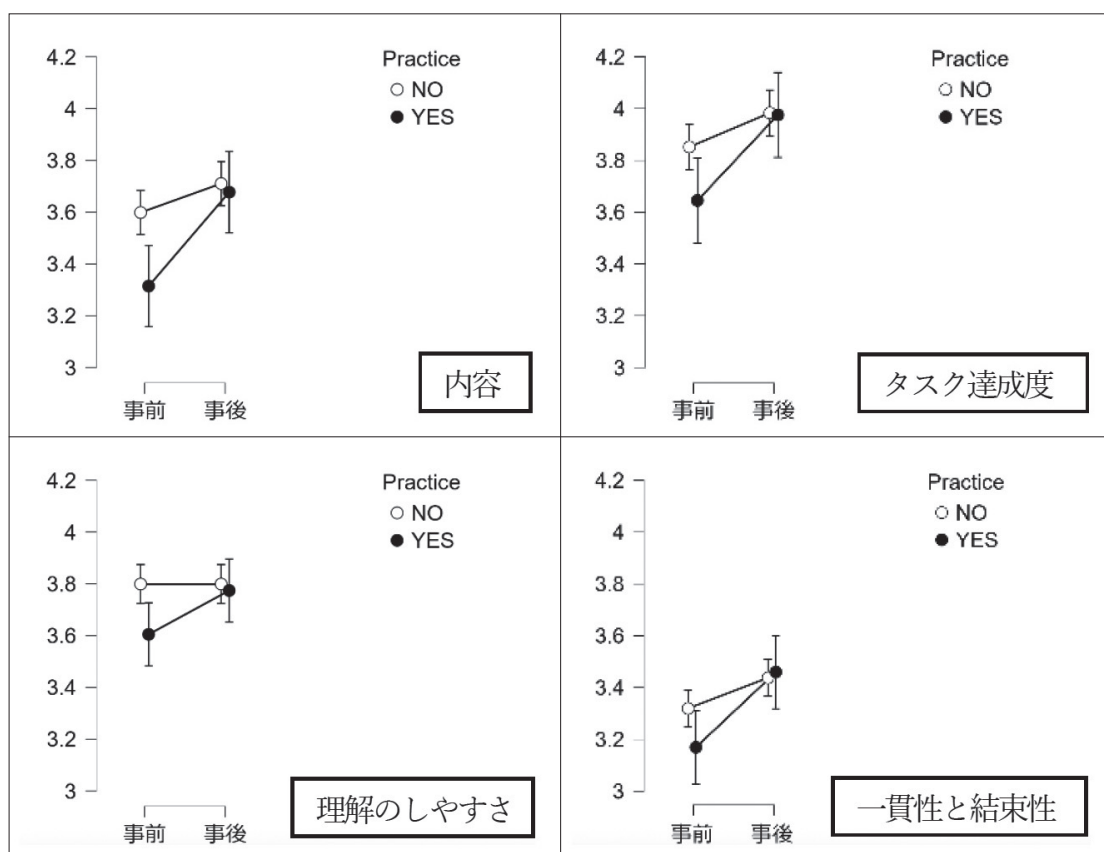
注: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

## 4.2 Gemini を活用した練習用課題の効果 (RQ2)

Gemini を活用した練習用課題が生徒の英語ライティングにおける機能的適切さに与える影響を検討するために、まず、機能的適切さの各観点ごとに、練習あり( $n = 62$ )と練習なし( $n = 152$ )のグループに分けて記述統計を確認した(表7, 図4)

■表7: グループ別の事前・事後テスト結果の記述統計

| 観点      | 時間 | 練習あり( $n = 62$ ) |      |       | 練習なし( $n = 152$ ) |      |       |
|---------|----|------------------|------|-------|-------------------|------|-------|
|         |    | $M$              | $SD$ | 事後－事前 | $M$               | $SD$ | 事後－事前 |
| 内容      | 事前 | 3.32             | 0.85 | 0.36  | 3.60              | 0.66 | 0.26  |
|         | 事後 | 3.68             | 0.65 |       | 3.86              | 0.53 |       |
| タスク達成度  | 事前 | 3.65             | 0.87 | 0.33  | 3.85              | 0.62 | 0.13  |
|         | 事後 | 3.98             | 0.66 |       | 3.98              | 0.59 |       |
| 理解のしやすさ | 事前 | 3.60             | 0.66 | 0.17  | 3.80              | 0.54 | 0.00  |
|         | 事後 | 3.77             | 0.50 |       | 3.80              | 0.51 |       |
| 一貫性と結束性 | 事前 | 3.17             | 0.73 | 0.29  | 3.32              | 0.62 | 0.12  |
|         | 事後 | 3.46             | 0.52 |       | 3.44              | 0.44 |       |



■図4: 機能的適切さの各観点におけるグループ別の事前・事後テスト結果の比較

両群とも平均はおおむね上昇したが、伸び幅は一貫して練習ありの方が大きかった。練習ありの事前と事後の差は、内容が0.36(練習なし0.26)、タスク達成度は0.33(練習なし0.13)、理解のしやすさは0.17(練習なし0.00)、一貫性と結束性は0.29(練習なし0.12)であった。

次に、練習の有無および事前・事後の差に対する統計的検定として、2要因の混合計画分散分析を実施し、時間の効果・練習の効果および交互作用の有無を検討した(表8)。

■表8: 観点別の2要因の混合計画分散分析の結果

| 観点     | 効果   | 項目        | 平方和    | df  | 平均平方 | F     | p        | $\eta^2p$ |
|--------|------|-----------|--------|-----|------|-------|----------|-----------|
| 内容     | 参加者内 | 時間        | 4.96   | 1   | 4.96 | 15.98 | <.001*** | .07       |
|        |      | 時間×練習     | 1.39   | 1   | 1.39 | 4.47  | .04*     | .02       |
|        |      | Residuals | 65.84  | 212 | 0.31 | —     | —        | —         |
|        | 参加者間 | 練習        | 2.22   | 1   | 2.22 | 4.25  | .04*     | .02       |
|        |      | Residuals | 110.61 | 212 | 0.52 | —     | —        | —         |
| タスク達成度 | 参加者内 | 時間        | 4.70   | 1   | 4.70 | 13.96 | <.001*** | .06       |
|        |      | 時間×練習     | 0.87   | 1   | 0.87 | 2.59  | .11      | .01       |
|        |      | Residuals | 71.42  | 212 | 0.34 | —     | —        | —         |
|        | 参加者間 | 練習        | 0.87   | 1   | 0.87 | 2.59  | .11      | .01       |

| 観点      | 効果   | 項目        | 平方和    | df  | 平均平方 | F     | p        | $\eta^2p$ |
|---------|------|-----------|--------|-----|------|-------|----------|-----------|
| タスク達成度  | 参加者間 | 練習        | 1.01   | 1   | 1.01 | 1.93  | .17      | .01       |
|         |      | Residuals | 111.62 | 212 | 0.53 | —     | —        | —         |
| 理解のしやすさ | 参加者内 | 時間        | 0.63   | 1   | 0.63 | 2.83  | .09      | .01       |
|         |      | 時間×練習     | 0.63   | 1   | 0.63 | 2.83  | .09      | .01       |
|         |      | Residuals | 47.24  | 212 | 0.22 | —     | —        | —         |
|         | 参加者間 | 練習        | 1.06   | 1   | 1.06 | 2.91  | .09      | .01       |
|         |      | Residuals | 77.43  | 212 | 0.37 | —     | —        | —         |
|         |      |           |        |     |      |       |          |           |
| 一貫性と結束性 | 参加者内 | 時間        | 3.68   | 1   | 3.68 | 16.06 | <.001*** | .07       |
|         |      | 時間×練習     | 0.65   | 1   | 0.65 | 2.84  | .93      | .01       |
|         |      | Residuals | 48.57  | 212 | 0.23 | —     | —        | —         |
|         | 参加者間 | 練習        | 0.36   | 1   | 0.36 | 0.87  | .35      | .00       |
|         |      | Residuals | 87.48  | 212 | 0.41 | —     | —        | —         |
|         |      |           |        |     |      |       |          |           |

注: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ 

結果として、交互作用が有意だったのは「内容」のみであった(時間×練習: $F = 4.47$ ,  $p = .04$ ,  $\eta^2p = .02$ )。その他の観点では交互作用は有意でなかった(タスク達成度: $F = 2.59$ ,  $p = .11$ ,  $\eta^2p = .01$ ; 理解のしやすさ: $F = 2.83$ ,  $p = .09$ ,  $\eta^2p = .01$ ; 一貫性と結束性: $F = 2.84$ ,  $p = .93$ ,  $\eta^2p = .01$ )。時間の効果は、内容( $F = 15.98$ ,  $p < .001$ ,  $\eta^2p = .07$ )、タスク達成度( $F = 13.96$ ,  $p < .001$ ,  $\eta^2p = .06$ )、一貫性と結束性( $F = 16.06$ ,  $p < .001$ ,  $\eta^2p = .07$ )で有意、理解のしやすさでは有意ではなかった( $F = 2.83$ ,  $p = .09$ ,  $\eta^2p = .01$ )。練習の効果は、内容のみ有意( $F = 4.25$ ,  $p = .04$ ,  $\eta^2p = .02$ )で、その他は有意でなかった。

交互作用が有意であった内容について単純主効果をみると(表9)、時間の効果は練習ありで有意( $F = 10.69$ ,  $p < .01$ )だが、練習なしでは有意ではなかった( $F = 3.37$ ,  $p = .09$ )。また、グループ間差は、事前で練習なし>練習ありで有意差が見られた( $F = 6.91$ ,  $p < .01$ )が、事後では有意差なしであった。( $F = 0.15$ ,  $p = .70$ )。練習あり群は事後に得点が伸びたことで、事前に見られた有意差が解消されたことがわかる。

まとめると、Geminiを活用した練習課題は、機能的適切さにおける内容の観点に対して効果があった。他の観点に対しては、全体として時間経過による向上は確認できたものの、統計的に有意な差は確認できず、限定的であったといえる。

■表9: 内容の観点における単純主効果

| 時間   |      |    |      |       |        |
|------|------|----|------|-------|--------|
|      | 平方和  | df | 平均平方 | F     | p      |
| 練習あり | 4.08 | 1  | 4.08 | 10.69 | <.01** |
| 練習なし | 0.95 | 1  | 0.95 | 3.37  | .09    |
| 練習   |      |    |      |       |        |
|      | 平方和  | df | 平均平方 | F     | p      |
| 事前   | 3.56 | 1  | 3.56 | 6.91  | <.01** |
| 事後   | 0.05 | 1  | 0.05 | 0.15  | .70    |

注: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

### 4.3 学習者の Gemini からのフィードバックに対する認知と応答 (RQ3)

Gemini を活用した英語ライティングの練習課題に関するアンケートは146件の回答を得た(回収率68%, 146/214)。なお, 練習なしのグループもアンケートに回答しており, その結果も含んでいる。

#### 4.3.1 肯定的なコメント(76%, 111/146)

全体として, 生徒は AI を活用した学習について, 全体的に非常に肯定的な意見を寄せていた。特に, 以下の点が特筆すべき傾向として見られる (表10)。

■表10: アンケートの肯定的なコメント

| 分類                     | コメントの抜粋                                                                                                                              | 分析                                                                                                                |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| 効率性と即時性<br>(25%)       | 「すぐに応答が返ってくるので, 時短になる」<br>「集中力が続かないときでもすぐに使える」<br>「いつでもどこでも添削してもらえる」                                                                 | AI の即時的なフィードバックと迅速な対応を高く評価していた。従来の教師による添削では時間を要した部分が, AI の活用により効率化され, 学習サイクルが速くなるというメリットを感じていることが窺える。             |
| 的確なフィードバックと学び<br>(24%) | 「自身の文法ミスやスペルミスを正確に指摘してくれた」<br>「より自然な表現や語彙を教えてくれた」<br>「なぜ間違っているのか, どうすればよいのかを具体的に示してくれた」                                              | AI の具体的なフィードバックで弱点・改善点が明確になり, 学習効果が高まり自律学習につながったとの声が多かった。さらに「新たな発見・気づき」が多く報告され, AI が新たな視点を与えていることが示された。           |
| ライティング力の向上<br>(15%)    | 「以前より文章が長く書けるようになった」<br>「話のつなげ方や, より適切な表現が深く理解できた」<br>「言いたいことをより正確に伝えられるようになった」<br>「構成を工夫するようになった」                                   | AI を活用した練習を繰り返すことで, 自身のライティング能力に手応えを感じている生徒もいた。機能的適切さの観点に関わる部分で改善できたというコメントもあったことから, AI によるフィードバックが活かされていることが窺える。 |
| 心理的な安心感<br>(12%)       | 「あまり先生に聞くことはなかったからいい体験だった」<br>「AI の添削で安心して学べた」<br>「気軽にアドバイスを求めることができるので, AI を参考にして英文を書くのもいいなと思った」<br>「スペルミスや文法の間違いを気にせずに作文をすることができた」 | 教師に直接質問することに抵抗を感じていた生徒にとって, AI は安心して学べる存在となっていたようである。些細な疑問でも, 気軽に AI に尋ねられる環境が提供されていることがわかる。                      |

### 4.3.2 批判的なコメント(24%, 35/146)

肯定的な意見が多数を占める一方で、批判的な意見も見られた(表11)。これらの意見は、英語ライティング指導に生成 AI を取り入れていく際に必要な学習者視点を提供している。

■表11: アンケートの批判的なコメント

| 分類                    | コメントの抜粋                                                                                                   | 分析                                                                                        |
|-----------------------|-----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| ツールとしての使いづらさ<br>(10%) | 「評価基準をコピーする作業が手間」<br>「パソコンで打つのが大変だったり、時間がかかったりする」<br>「手で書く方がやりやすい」                                        | AI とのやり取り、そもそも PC でのタイピングに慣れていないことによる不便さを訴える意見があった。継続的に使用することで、解消できるか。                    |
| 根本的な英語力の不足<br>(8%)    | 「そもそも書きたい単語のスペルがわからなかった」<br>「文法が曖昧で混乱する」<br>「話の広げ方が難しく、トピックによってはあまり書けなかった」<br>「なぜその表現が良いのか、理由がわからないことがある」 | 根本的な語彙力や文法力に課題を感じている生徒もいた。AI はあくまでツールであり、自身の基礎力がなければ最大限に活用することは難しいと、生徒が感じていることが窺える。       |
| AI への依存に対する懸念<br>(6%) | 「AI に頼りすぎると、それに頼った文章しか作れなくなる」<br>「このままだと人間である意味がなくなる」<br>「自身の能力を潰されている気がする」                               | AI の利便性からくる依存性や、学習効果に対する懸念も一部で見られた。これらの意見は、AI を何のためにどのように使うのか、という根本的な問題について考える必要性を示唆している。 |

### 4.3.3 フィードバックに基づいた修正

練習ありグループの中で、事前テストから事後テストにかけてライティングの評価が顕著に向上した2名の生徒の練習用課題におけるライティングについて、Gemini からのフィードバック(資料1)を受けてどのように修正が行われたかを分析した。一人目は、練習②の“Something I use every day”をトピックとして書いたものである。なお、生徒のライティングは、スペルミスや文と文の間のスペースが空いてないものなどもそのまま載せている。

#### 【修正前】

Hello! I'm going to show you what I use every day.It's the smartphone. In modern Japan, the smartphone has become an indispensable tool for life because it plays various roles such as a game, a stopwatch, a timer, an alarm clock, a clock, and a dictionary.Smartphones can also serve as a substitute for telephones.That's why I'm always on my phone.Thank you for reading!

#### 【修正後】※下線, 太字は筆者による。Gemini からのフィードバックを踏まえて変更された部分。

Hello! I'm going to introduce you to a tool I use all the time. It's a smartphone. Smartphones are essential to life in modern Japan because they allow you to play games, make phone calls, read e-books, and use the internet to do research. **I play a game called Blue Archives on my**

smartphone every day. This game is operated by a company called Yostar and is extremely popular. I also read light novels every day using an application called Kakuyomu. Light novels are stories with light content that are easy to read even for beginners. That's why I use my smartphone every day to play games and read e-books. Please tell me about the tools you use every day. Thank you for reading!

#### 【Gemini からの改善のためのフィードバックの概要】

- ① スマートフォンの機能を具体的にどのように使っているかを説明する(内容)
- ② 読者に対して直接語りかけるような表現(例: "Imagine a world without it!" や "You might be surprised how much it helps me..." など)を加える(タスク達成度)
- ③ 文を主語から始めるだけでなく、時折、副詞句や前置詞句を文頭に持ってくることで、リズムが生まれより自然な英語になる(理解のしやすさ)
- ④ indispensable だけでなく essential, vital, crucial など、文脈に応じて使い分けることで文章に深みが増す(理解のしやすさ)
- ⑤ 多機能性から「肌身離さず持っている」という結論に至るまでの論理的なつながりを、よりスムーズにする接続表現を検討(一貫性と結束性)
- ⑥ 具体例から抽象的な結論へ橋渡しとなるような文(例: All these features combined make my smartphone an indispensable companion for daily life.)の挿入を検討(一貫性と結束性)

この中で、修正する際に基になったと思われるフィードバックは①(内容)、②(タスク達成度)、④(理解のしやすさ)である。ゲームと小説という具体的なスマホの機能の使用が追加され、indispensable が essential に変更されている。さらに、読者に問いかけるような1文を追加している。この生徒は、フィードバックの内容を自分なりに解釈して自分の言葉で修正を加えており、Gemini が提案している文をそのまま使うことはしていない、という点が興味深い点である。

二人目も、練習②の "Something I use everyday" をトピックとして書いたものである。

#### 【修正前】

I'll introduce my eye mask.  
My brother gave me it about half years ago, and I wear it every night now.  
It's printed character's face who is a Waddle Dee.  
I love it very much now. It's so cute and help my good sleep.  
I can't imagine if I don't get it. How could I sleep without it?  
If you have not ever use eye mask, you should try it.

【修正後】※下線、太字は筆者による。Gemini からのフィードバックを踏まえて変更された部分。

I'll introduce my eye mask.  
My brother gave it to me about half a year ago. Since then, I wear it every night now.  
It's printed character's face who is a Waddle Dee. My brother and I like Kirby's Dream Land series. Especially, I like Waddle Dee. So, the face on it comforts me.  
I love it very much now because it helps my good sleep.  
Eye mask block rights, so I can focus on my sleep.  
I can't imagine life without it. How could I sleep?  
If you have not ever use eye mask and want to good sleep, you should try it.

### 【Gemini からの改善のためのフィードバックの概要】

- ①アイマスクの具体的な利点(なぜそれが良い睡眠につながるのか、など)やキャラクターの情報について、もう少し詳しく記述する(内容)
- ②「もしあなたが良い睡眠を探しているなら、ぜひアイマスクを試してみてください」のように、読者へのはたらきかけを意識する(タスク達成度)
- ③“gave me it”を“gave it to me”に、“half years”ではなく“half a year”、“I can't imagine if I don't get it”は、“I can't imagine not having it”や“I can't imagine life without it”とするとより自然である(理解のしやすさ)
- ④“and”よりも、時間の経過を示す副詞句(例:“Since then”または“Ever since then”)を使って、アイマスクをもらってから毎晩使っているという関係性をより明確にする(一貫性と結束性)
- ⑤理由を説明する際に“because”などの接続詞を使うと、より論理的なつながりになる。例:“I love it very much because it's so cute and helps me get good sleep.”(一貫性と結束性)
- ⑥“It's printed character's face who is a Waddle Dee.”の後に、“This design makes it even more special to me.”のように、補足的な情報を加えることで、前の文とのつながりを強めることができる。(一貫性と結束性)

この生徒は、Gemini からのフィードバックをすべて取り入れるように修正していた。③と④、⑤のように単語やフレーズレベルの修正に関しては、提案された案をそのまま採用している。しかし、その他の点に関しては、具体的な英語の修正案が提示されていなくとも、自分の言葉で表現をして修正を試みている。例えば、①については、“Eye mask block rights, so I can focus on my sleep.”や“My brother and I like Kirby's Dream Land series.”といったように、自身の持っている語彙や文法を駆使して表現を試みている。

このように、事前テストから事後テストにかけてライティングのパフォーマンスが向上した二人の生徒は、自分の言葉で表現をしようとする姿勢が共通していた。

## 5 考察

### 5.1 英語ライティング指導における生成 AI を活用した機能的適切さについてのフィードバック

本研究では、高等学校における英語ライティング指導の中で、生徒の自律的学習を促すため、機能的適切さという構成概念を用いた生成 AI によるフィードバックの効果を Gemini を用いて実証的に検証した。まず、教師評価と Gemini による機能的適切さの評価は中程度の一致であり、Gemini は教師よりも一貫して高めの評価をする傾向が見られた。以下は一人の生徒が書いた英文だが、教師の評価が(内容:1.33, タスク達成度:2, 理解のしやすさ:2.67, 一貫性と結束性:3)なのに対して、Gemini の評価は(内容:4, タスク達成度:4, 理解のしやすさ:4, 一貫性と結束性:4)であった。

My last year birthday, I got a gift which my mother gave me. I am using it everyday when I go to school. It is very convenient for me to go to school, because it is simple shape but it can have lot of things .

これは、Gemini が教師に比べ相対的に寛容な評価を行うことを意味しており、教師による最終的な

補正が必要であることを示している。英文によって、教師評価とAI評価が一致している観点も多くあるため、プロンプトを調整して一概に2つの評価を近づけようとするのは難しいように思われる。Xi(2023)は、AI評価が「ブラックボックス」であり、用いられているモデルやアルゴリズムが説明できない点は問題であるとしながらも、最終的な人間による評価の必要性和、対象がローステークスなものであるならば、評価が多少ブラックボックスでも問題ないだろうと述べている。つまり、生徒がライティングの練習用課題に取り組む際、形成的なフィードバックを得るためにGeminiを活用することは問題ないと考えられる。加えて、川村他(2019)でも述べられていたような教師の負担を軽減することにもつながるだろう。

練習用課題の効果に注目すると、全体としての時間の効果は内容・タスク達成度・一貫性と結束性の観点において有意であった(理解のしやすさについては非有意)。内容の観点では、練習×時間の交互作用が有意で、練習ありグループのみ、事前から事後へ評価が有意に向上し(Holm 補正後も有意)、練習なしグループの向上は有意ではなかった。さらに、事前では練習なしグループの評価が有意に高かったが、事後ではその差は消失した。したがって、Geminiを活用した練習用課題は、内容の観点に限って効果を示し、アイデアの追加や充実を促した可能性が高い。一方、タスク達成度と一貫性・結束性は時間の効果のみで、練習の主効果・交互作用はいずれも有意ではなかった。これらの観点は、タスクの文脈に応じて内容や表現を調整する力や、接続表現や論理展開に関わる高度な文章構成力を要するため、今回の短期間・自律型の練習ではAIフィードバックが十分に運用されにくかったと考えられる。先行研究でも、自動フィードバックは文法や語彙といった表層的側面の修正に偏りやすいことが指摘されており(Liu & Yu, 2022)、本研究の結果も同様の傾向を示したと考えられる。加えて、理解のしやすさは時間の効果も練習の効果も確認できず、この観点については教師による明示的指導や支援といったような足場かけが必要であることが示唆される。事後テストにおいてパフォーマンスが向上した二人の生徒のライティングの分析からは、Geminiからのフィードバックを参考にして、自分の伝えたい内容を自分の言葉で表現しようとする姿勢が見られた。AIから提案された修正案を、一旦自分の伝えたい内容と照らして考えてから書く、ということが必要だと考えられる。

実際の数字には表れなかったものの、アンケート結果において、生徒たちは生成AIが英語ライティングの学習において非常に有効なツールであり、特に効率的なフィードバックや学習へのモチベーション向上に大きく貢献している、と感じていたことがわかった。Geminiからのフィードバックに対して全体の76%が肯定的に受け止めており、効率性や即時性、心理的な安心感が高く評価されていた。特に、「AIに直されたところが新たな発見になった」との記述は、水本(2025)でも述べられているように、生成AIによるフィードバックが言語習得における学習者の気づきを促していることを示唆している。一方で、生成AIからのフィードバックから言語的な気づきを得るためには、生徒自身の基礎的な英語力が不可欠である可能性も見えてきた。プロンプトの中で具体的なフィードバックをするように指示をしても、機能的適切さに関するフィードバックは抽象的な指摘にとどまる場合も少なくない。山下(2024)も、学習者が自ら修正につなげられないタイプのフィードバックがあることを指摘している。そのような場合、生徒自ら言語的な気づきを得られる程度の基礎力がなければ、フィードバックを活かすことができない。実際、Lu and Suzukida(2025)でも、熟達度の低い学習者に対しては、機能的適切さについてではなく、語彙の正確性に焦点を当てた指導が効果的である可能性を指摘し、習熟度に応じた支援の必要性を述べている。つまり、生成AIを用いた自律的な学習ができるようになるまでの足場かけは、教師が必ず担い、指導・支援すべき部分であるといえる。また、生徒からの「便利すぎて人間の意味がなくなりそう」との意見は、生成AIへの過度な依存への懸念を反映しているといえる。生成AIを使うことを目的にせず、生成AIを使って自身の能力を高めていくという姿勢が重要であるというような指導も、生成AIを英語教育に取り入れる際には必要だろう。

## 5.2 教育的示唆

本研究の結果から、英語ライティング指導における機能的適切さに関する生成 AI の活用について、複数の教育的示唆を得られた。

第一に、生成 AI によるフィードバックの活用は教師の役割を不要にするのではなく、むしろ役割を明確にする、ということである。生成 AI の機能的適切さに関する評価は、教師の評価と一定の一致率は認められたものの、一貫性があるわけではなく、また学習者がその内容を必ず理解できるとは限らないということが、本研究でも明らかになった。そのため、教師は生成 AI からのフィードバックを鵜呑みにさせず、学習者が批判的に受け止め、「なぜこの表現が良いのか」を生徒自身が考えるような場を設定し、自らの表現に適用できるような思考力が育まれる授業をデザインすることが重要だといえる。Wang and Wang (2025) は、L2 学習者のライティングにおける批判的 AI リテラシー (Critical AI Literacy; CAIL) が「批判的気づき」「立場性」「対話方略」「アフォードانسの評価」という4つの要素から構成されているとし、それらを高めていくためには、プロンプト設計・出力の検証・AI と人間フィードバックの比較検討を授業に組み込むことで、CAIL を「追加スキル」としてではなく、従来のライティング教育に統合していく必要性を指摘している。具体的な活動としては、生成 AI に入力したプロンプトと出力されたフィードバックをクラス全体で共有し、それを基にどのように修正をするのが適切かを議論させるような活動や、学習者がお互いのライティングに対するフィードバックを作成し、生成 AI からのフィードバックと比較するようなタスクを行うことで、学習者の「気づき (noticing)」を促進できると思われる。さらに、生成 AI からのフィードバックが抽象的で理解が難しいときには、AI と対話することで理解を深めるような練習を行う場をデザインすることも考えられる。本研究においては、水本 (2025) の ChatGPT を学習者に使用させる際の手順のうち、「必要に応じて ChatGPT に追加質問」という部分についてはあまり強調しなかったため、その点に関しても授業内で丁寧に指導してから自律的学習につなげるなど時間をかけて行うことで、ライティングのパフォーマンスへの効果は違ったものになる可能性もある。このように、生成 AI をあくまでも1つの学習ツールとして明確に位置づけ、教材として扱うことで、生徒の CAIL 向上にもつながると考えられる。そして、そのような学びを教室内で臨機応変にファシリテートする役割こそ、教師が担うべきものとなるだろう。

第二に、生成 AI によるフィードバックは即時性と豊富な出力を特徴とし、学習者にとって「安全な練習の場」を提供できる点が挙げられる。アンケートにおいても学習者の多くが肯定的に評価しており、心理的負担を軽減しながら学習を継続できる可能性が示された。従来、教師による個別フィードバックは時間的・労力的に制約が大きく、十分な量を提供することが難しかった。しかし、本研究において、教師による機能的適切さに関する評価との中程度の相関を示すことがわかったため、生成 AI はその補完的役割を担うことが可能だろう。特に「内容」の改善に寄与したという結果は、学習者が新たなアイデアを得たり発想を広げたりすることができるという点で、教育的意義が大きいといえる。逆に、その他の観点に関しては、本研究における自律的学習では有意な向上は見られなかったため、練習用課題の実施前や実施中に、より丁寧な説明や指導を行い、3つの観点に関しても向上につながるようにする手立てが必要であるといえる。アンケートの中で、「あまり先生に聞くことはなかったからいい体験だった」と答えていた生徒がいたように、教師に質問するのが苦手な生徒もいる。そのような生徒たちが生成 AI に気軽に質問して学んでいけるように、教師が授業の中で効果的な活用法を積極的に指導していくことが求められるだろう。

第三に、AI の利用には過度の依存が伴う危険性がある点に留意すべきである。本研究においても「AI に頼りすぎると、それに頼った文章しか作れなくなる」という生徒の声があった。AI への過度な依存は学習者の自律的な学習態度を損なう危険があるということを、自認してコメントしていた生徒がいたことは大変興味深い。全員がそれを自ら感じているわけではないだろう。高校の教育現場では、生成 AI を使わずに書く活動を明確にゴールタスクに据えることで、生成 AI の活用が生成 AI を使わずに書くことがで

きるようになるための手段と捉えられるだろう。單元ごとの指導で様々なライティングを扱い、それぞれに対する生成AIを活用した練習を授業内外で設計して教師が指導していくことで、生徒自身の英語力伸長のための生成AIの補助的な活用方法について学んでいくことができる。機能的適切さを評価することの教育的意義も大きい。誤りが残っていても、目的達成に十分であれば高く評価されるという視点は、学習者の自信を高め、実用的な言語運用能力の育成につながる。AIフィードバックを機能的適切さの観点と組み合わせて導入することは、学習者に「何が伝わる文章なのか」を理解させる手助けとなり、単なる文法的正確さの追求にとどまらない実践的なライティング指導を可能にするだろう。

### 5.3 研究の限界と今後の展望

本研究における限界と今後考えられる研究について述べる。まず、対象となった学習者が限定的であった点である。本研究では特定の高校に通う生徒を対象としたため、学習者の英語力や学習背景が比較的均質であった。したがって、本研究で得られた知見を異なる学力層や多様な学習背景をもつ学習者へ一般化するには注意が必要である。例えば、英語力が初級段階の学習者にとって生成AIの機能的適切さに関するフィードバックが過剰な負担となる可能性や、逆に上級者にとっては既知の情報にとどまり十分な学習効果を生まない可能性もあり、さらには中学生や大学生にとっては効果が変わってくる可能性もある。今後の研究として、異なる熟達度の学習者における調査をすることで、より機能的適切さに関する生成AIのフィードバックの効果を明らかにしていくことができるだろう。また、参加者のアンケートの中に、フィードバックの量が多すぎるという内容のコメントが複数見られたため、フィードバックの個数を限定するようにプロンプトを調整して実施することも考えられる。あるいは、生成AIからのフィードバックを読み取る際のガイドラインを作成し、着目するポイントを示すことも効果的かもしれない。

また、データ収集の方法と量に制約があった点も限界点として挙げられる。本研究では、実践が授業外での実施にとどまったため、長期的な学習成果や学習者の態度変容を追跡することはできなかった。そのため、生成AIからのフィードバックが持続的にライティング能力の向上に寄与するのか、あるいは一時的な効果にとどまるのかは今後の検討課題である。生徒からのアンケートの記述の一部からは、「構成を工夫するようになった」や「話のつなげ方や、より適切な表現が深く理解できた」など、機能的適切さに関する学びが起こっていたことが窺える。より長期的な縦断的研究や実験的デザインを取り入れることで、機能的適切さに関する生成AI活用についての中長期的効果をより明確に検証できるだろう。今後の研究として、間隔を空けて行ったり、学期中でなく長期休みの課題として行うことなども考えられる。今回は学期中で、他の教科の学習も忙しい中で行うには負担が大きすぎた可能性がある。また、事前テスト後に、1回目の練習は教師の指導の下、授業中に取り組ませるなど、手順が明確に伝わった状態で自律的学習を始める方が取り組みやすい可能性がある。加えて、実践的な授業内での研究もまた必要だろう。教師がAIフィードバックを授業の中でどのように組み込むのか、学習者がそれをどのように受け止め、自律的学習へとつなげるのかを具体的に検証することで、生成AIと人間の協働による新たなライティング教育のモデルを構築できる可能性がある。

ルーブリックや採点に関しても、さらなる研究の余地がある。今回は、Kuiken and Vedder(2017)の機能的適切さのルーブリックのみを使用した。語彙や文法といったような言語形式に関するルーブリックも加えて同時に評価することで、学習者のライティングパフォーマンスの伸びをより明確に把握することができるだろう。本研究では生成AIによる採点を行う際に、プロンプトに入力した採点例が6つのみであり、1から6の採点が包括的に入るような例になっていなかった。また、AIには採点スコアのみ提示し、なぜその採点になるかという説明をプロンプトに入れていなかった。プロンプトに採点基準についてのより詳細な情報を入力することで、AIによる評価を教師による評価により近づけられる可能性もあるかもしれない。

さらに、生成AIからのフィードバックの質に関する分析も不十分である。本研究では機能的適切さの

観点を導入したが、AIのフィードバックそのものの質や特徴についての検証や、それらに対して学習者がどのように受け止め修正に活かすのか、修正につながりやすいものとつながりにくいものの違いなどについて質的に詳らかにできれば、それらは生成AIから適切なフィードバックを出力させるためのプロンプト改善につながるだろう。また、異なるAIモデルを比較し、どのモデルがどのような観点で有効性を示すかを明らかにすることや、生成AIと教師のフィードバックを組み合わせる最適な方法を探ることも考えられる。

## 6

## おわりに

本研究は、高等学校の英語ライティング指導において、機能的適切さに関する生成AIを用いたフィードバックの効果を検証した。21日間という限られた期間内での実施ではあったが、本研究で行った実践が生徒のライティングのパフォーマンスのうち、機能的適切さの内容に関わる部分の向上に寄与しているということが確認できた。また、教師の評価と生成AIの評価がある程度一致するということや、生徒は機能的適切さに基づいた生成AIからのフィードバックに関して肯定的に捉え、学びにつながるという実感をもっていたことがわかった。教師が行っていたライティング指導の一部を生成AIに任せられる可能性や、その際に使用できる評価の基準として、機能的適切さの有用性も見えてきた。本研究が、生徒が生成AIを活用して自律的に英語ライティングを学ぶための指導設計、およびその関連の研究の一助となれば幸いである。

### 謝辞

本研究を実行する機会を与えてくださった公益財団法人 日本英語検定協会、関係者の皆様、ならびに、選考委員の先生方に、心より御礼申し上げます。実施に当たって、小泉利恵先生からは、大変お忙しい中、本研究に関する多くのご助言、アイデアを賜りました。また、研究依頼校の先生方や生徒の皆さんからは、熱心な取り組みと貴重な振り返りのコメントを頂きました。加えて、2名の採点協力者は真摯に、そして迅速に採点に取り組んでくださいました。研究にご協力くださった方々に心よりお礼申し上げます。なお、本報告書に関して、開示すべき利益相反関連事項はありません。

## 引用文献

- Allen, T. J., & Mizumoto, A. (2024). ChatGPT over my friends: Japanese English-as-a-foreign-language learners' preferences for editing and proofreading strategies. *RELJ Journal*, 0(0). <https://doi.org/10.1177/00336882241262533>
- De Jong, N., Steinel, M., Florijn, A., Schoonen, R., & Hulstijn, J. (2007). The effect of task complexity on fluency and functional adequacy of speaking performance. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, I. Vedder, & K. Contactforum (Eds.), *Complexity, Accuracy and Fluency in Second Language Use, Learning & Teaching* (pp. 53-63). KVAB, Universa Press, Wetteren.
- Feng, H., Li, K., & Zhang, L. J. (2025). What does AI bring to second language writing? A systematic review (2014-2024). *Language Learning & Technology*, 29(1), 1-27. <https://doi.org/10.64152/10125/73629>
- Huang, J., & Mizumoto, A. (2025). The role of generative AI in mediating L2MSS and engagement with written feedback in EFL learning: A structural equation modeling approach. *Annual Review of Applied Linguistics*, 1-17. <https://doi.org/10.1017/S0267190525000029>
- 門田修平 (2024).『AIフル活用! 英語発信力トレーニング』コスモピア.
- Karatay, Y., & Karatay, L. (2024). Automated writing evaluation use in second language classrooms: A research synthesis. *System*, 123, Article 103332. <https://doi.org/10.1016/j.system.2024.103332>
- 川村晃市・柏木治美・康敏 (2019).「教師の英作文添削を支援する誤り検出システム」『日本教育工学会研究報告集』19(1), 83-88. <https://hdl.handle.net/20.500.14094/90005737>
- Koizumi, R., & In'nami, Y. (2022). Assessing functional adequacy using picture description tasks in classroom-based L2 speaking assessment. *JLTA Journal*, 25, 60-79. [https://doi.org/10.20622/jltajournal.25.0\\_60](https://doi.org/10.20622/jltajournal.25.0_60)
- Koizumi, R., & In'nami, Y. (2024). Predicting functional adequacy from complexity, accuracy, and fluency of second-language picture-prompted speaking. *System*, 120, Article 103208. <https://doi.org/10.1016/j.system.2023.103208>
- Kuiken, F., & Vedder, I. (2017). Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing*, 34(3), 321-336. <https://doi.org/10.1177/0265532216663991>
- Liu, S., & Yu, G. (2022). L2 learners' engagement with automated feedback: An eye-tracking study. *Language Learning & Technology*, 26(2), 78-105. <https://doi.org/10.125/73480>
- Lu, X., & Suzukida, Y. (2025). Relationships of linguistic complexity and accuracy to functional adequacy across L2 English speaking and writing: Roles of L2 proficiency and task type. *International Journal of Applied Linguistics*. Advance online publication. <https://doi.org/10.1111/ijal.12775>
- 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), Article 100050. <https://doi.org/10.1016/j.rmal.2023.100050>
- 水本 篤 (2025).「AIとライティング教育—英語ライティング教育における生成AIの活用と課題—」. 李 在鎬・青山玲二郎(編著)『AIで言語教育は終わるのか? 深まる外国語の教え方と学び方』(第8章) くろしお出版.
- 文部科学省 (2018).『高等学校学習指導要領(平成30年告示)解説 外国語編 英語編』[https://www.mext.go.jp/a\\_menu/shotou/new-cs/1407074.htm](https://www.mext.go.jp/a_menu/shotou/new-cs/1407074.htm)
- 文部科学省 (2024).『初等中等教育段階における生成AIの利活用に関するガイドライン Ver. 2.0』[https://www.mext.go.jp/a\\_menu/other/mext\\_02412.html](https://www.mext.go.jp/a_menu/other/mext_02412.html)
- Révész, A., Ekiert, M., & Torgersen, E. N. (2016). The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied Linguistics*, 37(6), 828-848. <https://doi.org/10.1093/applin/amu069>
- Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, 36(2), 187-209. <https://doi.org/10.1017/S0958344023000265>
- 鈴木駿吾 (2023).「第二言語スピーキング能力の機能的達成度は自動判定できるか?」『日本音響学会誌』79(3), 177-183. [https://doi.org/10.20697/jasj.79.3\\_177](https://doi.org/10.20697/jasj.79.3_177)
- Tate, T. P., Steiss, J., Bailey, D., Graham, S., Moon, Y., Ritchie, D., Tseng, W., & Warschauer, M. (2024). Can AI provide useful holistic essay scoring? *Computers and Education: Artificial Intelligence*, 7, Article 100255. <https://doi.org/10.1016/j.caeai.2024.100255>
- Voss, E., Cushing, I., Ockey, G., & Yan, D. (2023). The use of assistive technologies including generative AI by test takers in language assessment: A debate of theory and practice. *Language Assessment Quarterly*, 20(4-5), 514-533. <https://doi.org/10.1080/15434303.2023.2273799>
- Xi, X. (2023). Advancing language assessment with AI and ML—leaning into AI is inevitable, but can theory keep up? *Language Assessment Quarterly*, 20(4-5), 357-376. <https://doi.org/10.1080/15434303.2023.2291488>
- Wang, C., & Wang, Z. (2025). Investigating L2 writers' critical AI literacy in AI-assisted writing: An APSE model. *Journal of Second Language Writing*, 67, 101187. <https://doi.org/10.1016/j.jslw.2025.101187>
- 山下美朋 (2024).「激震の英語ライティング教育」山中司 [編] (2024)『AI・機械翻訳と英語学習 教育実践から見えてきた未来』(pp.35-57) 朝日出版社.
- 山下美朋・山中司・杉山滉平 (2024).「AIツールを活かした英語ライティング授業—英文作成支援ツールTransableを導入して—」『立命館高等教育研究』24, 75-87. <https://ritsume.repo.nii.ac.jp/records/2000853>

資料1-1: Geminiからのフィードバックの例(一人目).....  
(事前テストから事後テストにかけてライティングの評価が顕著に向上した2名)

|       | 内容   | タスク達成度 | 理解のしやすさ | 一貫性と結束性 |
|-------|------|--------|---------|---------|
| 事前テスト | 2.67 | 2.67   | 3       | 2.33    |
| 事後テスト | 4.67 | 4.67   | 4.67    | 4.33    |

評価と具体的な改善アドバイス

内容 (Content)

評価: 3

アイデアの数はある程度適切ですが、あまり一貫性がありません。スマートフォンが「日用品」であること、そしてその多機能性については述べられていますが、具体的な「あなたが」どのように使っているかという情報が不足しており、読者に紹介するという目的からすると、個人的な視点やエピソードがほとんどありません。

改善のためのアドバイス:

具体的な使用例の追加: スマートフォンが「ゲーム、ストップウォッチ、タイマー、目覚まし時計、時計、辞書」として機能するとありますが、あなたがそれらを具体的にどのように使っているのかを説明しましょう。「ゲーム」を例にとるなら、「私は朝の通学中に〇〇というゲームをして気分転換をしています」のように、具体的なゲーム名や、なぜそのゲームをするのか、それがあなたにとってどのような意味を持つのかを付け加えると良いでしょう。

個人的なエピソードの導入: 「なぜそれがあなたにとって欠かせないのか」をより深く掘り下げてください。例えば、「一度スマートフォンを家に忘れて出かけた時、どれほど不便を感じたか」といった個人的な経験を短く加えることで、読者はあなたの主張に共感しやすくなります。

タスク達成度 (Task requirements)

評価: 4

「Something I Use Every Day」というトピックに対して、スマートフォンを選び、その多機能性について述べているため、大半の要件は満たしています。しかし、「他人に紹介する内容」という点では、個人的なつながりや魅力が十分に伝わっていません。

改善のためのアドバイス:

「紹介」の視点を意識する: 読者が「へえ、そんな使い方もあるんだ!」と思ったり、「私も使ってみようかな」と感じたりするような「紹介」の要素を強化しましょう。単に機能の羅列ではなく、その機能があなたの日常生活をどのように豊かにしているのか、便利にしているのかを具体的に伝えることが重要です。

読者への語りかけ: 読者に対して直接語りかけるような表現(例: "Imagine a world without it!" や "You might be surprised how much it helps me..." など)を少し加えることで、「紹介」のニュアンスが強まります。

(資料1-1 続き).....

理解のしやすさ (Comprehensibility)

評価：5

文章は非常に理解しやすく、読みやすいです。文章構造がシンプルで、語彙も難しくないため、一読で内容を理解できます。

改善のためのアドバイス：

表現の多様化：理解しやすいことは素晴らしいですが、さらに読者の興味を引きつけるために、少し表現の幅を広げることを検討しても良いでしょう。例えば、同じ「とても大切」を意味する言葉でも、"indispensable" だけでなく "essential", "vital", "crucial" など、文脈に応じて使い分けることで、文章に深みが増します。

文頭のバリエーション：すべての文が「It's the smartphone.」「In modern Japan...」「Smartphones can also...」「That's why...」のように、主語から始まることが多いです。時折、副詞句や前置詞句を文頭に持つてくることで、リズムが生まれ、より自然な英語に近づきます。

一貫性と結束性 (Coherence and cohesion)

評価：4

文章には一貫性があり、無関係な文の流れは稀です。スマートフォンという共通のトピックを中心に記述されており、接続詞も良好に使われています。

改善のためのアドバイス：

接続表現の洗練：「That's why I'm always on my phone.」という締めくくりの文は、それまでの多機能性の羅列に対する結論としては少し唐突に感じるかもしれません。多機能性から「肌身離さず持っている」という結論に至るまでの論理的なつながりを、よりスムーズにする接続表現を検討しましょう。

例：「These diverse functions make my smartphone an extension of myself, so I truly can't imagine a day without it.」のように、多機能性が「肌身離せない」理由であることをより明確に示すと良いでしょう。  
具体例と抽象化のバランス：「ゲーム、ストップウォッチ ...」と具体例が列挙された後、「That's why I'm always on my phone.」と抽象的な結論につながっています。この間のギャップを埋めるために、具体例から抽象的な結論へ橋渡しとなるような文（例：「All these features combined make my smartphone an indispensable companion for daily life.」）を挿入することを検討すると、より結束性が高まります。

資料1-2: Geminiからのフィードバックの例(二人目) .....  
(事前テストから事後テストにかけてライティングの評価が顕著に向上した2名)

|       | 内容   | タスク達成度 | 理解のしやすさ | 一貫性と結束性 |
|-------|------|--------|---------|---------|
| 事前テスト | 3.67 | 3.33   | 3       | 3.67    |
| 事後テスト | 5    | 5.33   | 4.67    | 5       |

「Something I Use Every Day」というトピックで書かれた英文、拝見しました。よく書けていますね。  
あなたのアイマスクへの愛情が伝わってきます。

それでは、ループリックに沿って、各観点から評価と具体的な改善アドバイスをさせていただきます。

内容 (Content)

評価：4

アイデアの数は適切で、アイマスクを紹介するという点において十分に一貫しています。アイマスクをもらった経緯、キャラクターのデザイン、そしてそれが睡眠に役立っていることなど、必要な情報が盛り込まれています。

改善のためのアドバイス：

アイマスクの具体的な利点について、もう少し詳しく記述すると、読者にとってさらに興味深い内容になります。

例えば、\*\*「なぜそれが『良い睡眠』につながるのか」\*\*について、具体的に説明を加えてみましょう。「光を完全に遮断してくれるから」「目の疲れが取れる感じがするから」など、具体的な感覚や効果を記述することで、読者はその利点をより鮮明に想像できます。

また、「Waddle Dee」のキャラクターについて、もう少し情報を加えることもできます。例えば、「Waddle Dee は有名なゲームのキャラクターで、そのとぼけた表情がかわいくて癒される」といった補足があれば、Waddle Dee を知らない読者にもその魅力を伝えられます。

タスク達成度 (Task requirements)

評価：5

「他人に紹介する内容」という課題の要件をほとんど満たしています。自分の使っているもの(アイマスク)を紹介し、その理由や効果、そして読者への勧めまで書かれています。

改善のためのアドバイス：

全体的に課題は達成できていますが、読者への働きかけをもう少し意識すると、さらに魅力的な紹介文になります。

例えば、「How could I sleep without it?」という問いかけは良いですが、その後に読者に具体的な行動を促すような一文を加えることもできます。「もしあなたが良い睡眠を探しているなら、ぜひアイマスクを試してみてください」のように、読者に直接語りかける形にしてみましょう。

また、「おすすめポイント」を箇条書きでまとめるなど、紹介文としての体裁をさらに整えることも検討してみてください。

(資料1-2 続き).....

理解のしやすさ (Comprehensibility)

評価：5

文章は非常に理解しやすく、読みやすいです。書かれている内容も明確で、読者が文章の目的やアイデアを理解する上で、特に努力を必要としません。

改善のためのアドバイス：

ほとんど問題ありませんが、さらに自然な英語表現にすることで、より滑らかに読めるようになります。

例えば、「My brother gave me it about half years ago」は、「My brother gave it to me about half a year ago」とする方が自然です。「half years」ではなく「half a year」が正しい表現です。

「I can't imagine if I don't get it」は、「I can't imagine not having it」や「I can't imagine life without it」とすると、より自然な英語表現になります。「get it」は「手に入れる」という意味合いが強く、この文脈には少し不自然です。

一貫性と結束性 (Coherence and cohesion)

評価：4

文章には一貫性があり、アイマスクについての内容で統一されています。接続詞もいくつか使われており、アイデア間のつながりは概ね良好です。

改善のためのアドバイス：

アイデア間のつながりをより明確にするために、接続詞や接続表現をさらに効果的に使うことで、文章全体の結束力を高めることができます。

文頭で\*\*「My brother gave me it about half years ago, and I wear it every night now.」\*\*のように、「and」で文をつなぐのは良いですが、時間の経過を示す副詞句（例："Since then, " または "Ever since then, "）を使って、アイマスクをもらってから毎晩使っているという関係性をより明確にすることができます。

「I love it very much now. It's so cute and help my good sleep.」のように、理由を説明する際に\*\*「because」\*\*などの接続詞を使うと、より論理的なつながりになります。例：「I love it very much because it's so cute and helps me get good sleep.」

文と文の接続に変化を持たせることで、単調さを避け、より洗練された印象を与えることができます。例えば、「It's printed character's face who is a Waddle Dee.」の後に、「This design makes it even more special to me.」のように、補足的な情報を加えることで、前の文との繋がりを強めることができます。

このフィードバックが、あなたのライティングスキル向上の一助となれば幸いです。もし他に何か質問があれば、遠慮なく聞いてくださいね

## 第37回 研究助成

## C 調査部門・報告 I ・英語教育関連の調査・アンケートの実施と分析

## 動機づけ方略に対する英語教員の認識と 実際の使用の関係についての考察

研究者: 川光 大介 大阪府／大阪公立大学工業高等専門学校 講師・大阪府／関西大学大学院 在籍

《研究助言者: 寺内 一》

## 概要

外国語(L2)学習者の動機づけを高めるためにL2教員が行う働きかけは、「動機づけ方略(Motivational Strategies; 以下, MS)」(Dörnyei, 2001)とよばれる。先行研究から, L2教員のMS使用には様々な要因が影響を及ぼすことや, その要因は教育環境によって異なることが指摘されている。本研究では, 近畿の工業高等専門学校(高専)で英語の授業を担当する教員を対象に2つの調査を行い, 彼らのMS使用に影響を与える要因を検討した。調査1では, MSに関する先行研究と同様, 当該高専の英語教員が特定のMSを効果的だと認識していても, その認識が必ずしも実際のMS使用につながっていないことが明らかになった。さらに調査2では, 英語教員のMS使用に影響を与える要因として, 先行研究で示されている要因に加え, 調査の対象とした高専特有のものと考えられる要因もあることが明らかになった。

### 1 はじめに

外国語(L2)学習に関して, 動機づけは常に議論の中心となってきた(Guilloteaux & Dörnyei, 2008; Lee & Pun, 2021)。これまでL2研究では, Gardner and Lambert(1972)の社会心理学的アプローチにおける統合的／道具的動機づけ(例: Gardner, 1985), Ryan and Deci(2000)の自己決定理論における内発的／外発的動機づけ(例: Agawa & Takeuchi, 2017; Noels et al, 2000; Yashima et al., 2009), Dörnyei(2009)のL2自己動機づけシステムにおける理想自己／義務自己(例: Fukui & Yashima, 2021; Taguchi, Magid & Papi, 2009; Ueki & Takeuchi, 2013), またDirected Motivational Currents(DMC)(例: Dörnyei, Ibrahim & Muir, 2015; Henry, Davydenko & Dörnyei, 2015)など様々な理論的枠組みが用いられてきた。これらの研究は, 研究目的によって異なる理論的枠組みを採用しているものの, 動機づけがL2学習において中心的な役割を果たすという点については広く一致した見解が得られている(Lee & Pun, 2021)。また, 学習者が学業的な課題に対して夢中になって取り組んでいる状態を指す「学習者エンゲージメント」(student engagement)(廣森・和田, 2024; Mercer & Dörnyei, 2020)においても, 動機づけは学習行動を生み出すために不可欠な要素として特に重要視されている。

他の学習者要因, 例えば学習スタイル(河合, 2010)や学習方略(尾関, 2010)と比べても, L2学習において, 動機づけは極めて重要な役割を担っている。動機づけは, 他の学習者要因と比べるとL2学習成果に対する説明率が高く, 研究によって異なるが, 15~25%程度の説明率があるとされている(竹内, 2023)。動機づけが学習者要因の中でも教育による可塑性が高い(竹内, 2010), すなわち教育的介入によって変化が期待できる学習者要因であることを考えると, 教育の場面では学習者の動機づけに働きかけること

が重要であると考えられる。

しかし、日本の学校において英語学習に積極的になれない学習者は多い(竹内, 2016)。平成29年度に高校3年生を対象に実施された文部科学省の調査(文部科学省, 2017)では、英語の学習が「好きではない」と回答した生徒が23.2%に上り、「どちらかといえば好きではない」と回答した生徒を含めると、その割合は52%に達する。この数値は、竹内(2016)が紹介している平成27年度の調査結果(文部科学省, 2015)における58.3%と比較するとやや減少しているものの、依然として半数以上の生徒が英語学習に否定的な態度を示しており、英語学習者の動機づけやその維持が急務であることを示していると考えられる。

学習者の動機づけに、教員は大きな影響力を持っている(Dörnyei, 2001; Dörnyei & Csizér, 1998; Dörnyei & Ushioda, 2011; 廣森, 2010)。L2学習者の動機づけを高めるために教員が行う働きかけは「動機づけ方略(Motivational Strategies: 以下, MS)(Dörnyei, 2001)」とよばれる。MSに関する先行研究は、L2教員によるMSの使用が学習者に対して有効であることを示している(Alrabai, 2016; Moskovsky et al., 2013; Lee et al., 2020; Papi & Abdollahzadeh, 2012; Sugita & Takeuchi, 2010; Sugita-McEown & Takeuchi, 2014; Wong, 2013)。しかし一方で、L2教員が効果的だと考えるMSを必ずしも授業で十分に実践していないことも報告されており(Cheng & Dörnyei, 2007; Dörnyei & Csizér, 1998; Guilloteaux, 2013; Ye & Hu, 2024)、L2授業に関連する様々な要因が教員のMS使用に影響を与えていることが示されている(例: Lee & Pun, 2021; Ye & Hu, 2024)。さらに、MSの使用に影響を与える要因は文脈によって大きく異なるため、こうした要因を理解するには、特定のL2学習環境に焦点を当てた調査が不可欠である(Ye & Hu, 2024)。本研究では、この観点から、日本の特定のEFL環境において英語教員のMS使用に影響を及ぼす要因を検討した。

## 2 MSについての先行研究

MSとは一般的に「学習者の動機づけを引き出し、刺激するために教員が行う教育的介入」(Guilloteaux & Dörnyei, 2008, p.57)を指す。より具体的には、教員が「学習者の動機づけを意識的に生み出し、高めるとともに、動機づけられた行動を継続して維持し、気が散る行動や競合する行動から学習者を守る」(Dörnyei & Ushioda, 2011, p. 103)ために用いる技術として定義されている。なお、Guilloteaux and Dörnyei(2008)では上記の定義に加え、「学習者が自分の動機づけを管理するために意識的に用いる自己調整ストラテジー(p. 57)」とも定義されているが、本論文では前者の側面、つまり教員が学習者の動機づけに働きかけるために用いる技術にのみ注目することとする。

MSの研究は、学習者の動機づけを引き出すために教員が用いるテクニックのリストを作成することから始まった。MS研究のパイオニアとして知られるDörnyei and Csizér(1998)は、ハンガリーのL2教員が学習者の動機づけを引き出すためにどのようなテクニックを使っているかを調べ、「動機づけに関する10の方法」(日本語訳は竹内, 2016)を明らかにした。その後、Dörnyei(2001)がDörnyei and Csizér (1998)を理論的に拡張し、35のマクロストラテジー、計102のMSで構成されるフレームワークを提唱している。Dörnyei(2001)のフレームワークは、その明瞭さと包括性、そしてDörnyei and Ottó(1998)が提唱したプロセスモデル(動機づけを動的なものとして捉え、時間軸の観点から動機づけ要因をモデル化したもの)との密接な関連性から、教室の現場で容易に用いることができるものとなっている(Lee & Pun, 2021)。

しかしながら、これらのMSの中には個別のL2学習環境を考慮した際、利用しにくい、または利用できないものも含まれている(竹内, 2016)。Sugita(2009)は、Dörnyei(2001)の102項目を日本のEFL環境に適合させるために65項目に整理した。さらに、Kawamitsu and Takeuchi(2024)は、工業高等専門学校を学生を対象にこの65項目を用いて探索的因子分析を行い、「学生主体の活動」、「自律のサポート」、「達成感を得られやすい授業展開」、「興味・関心の刺激」、「外発的動機づけの促進」の5因子構造を

明らかにしている。

MSに関する先行研究では主に、Dörnyei and Csizér(1998)やDörnyei(2001)によるMSのリストを用いて、L2教員によるMSが学習者の動機づけ(Guilloteaux & Dörnyei, 2008; Sugita, 2009; Sugita & Takeuchi, 2010; Sugita-McEown & Takeuchi, 2014)や教室内外での学習行動(Guilloteaux & Dörnyei, 2008; Lee et al., 2020),さらにはL2の達成度(Alrabai, 2016)に影響を与えることが示されてきた。また、一部のMSは使用頻度が低くても学生の動機づけに影響を与えることがあること(Sugita, 2009; Sugita & Takeuchi, 2010)や、学習者のL2習熟度や元の動機の強さ、そして教員がMSを使用するタイミングが、MSの効果を左右する重要な要因であることもわかっている(Sugita, 2009; Sugita & Takeuchi, 2010; Sugita-McEown & Takeuchi, 2014)。

一方、L2教員によるMSの効果の認識と実際のMS使用の関係は、必ずしも直線的ではないことが明らかになっている。Cheng and Dörnyei(2007), Dörnyei and Csizér(1998), Guilloteaux(2013),そしてYe and Hu(2024)は、L2教員から得られたMSに関する回答をzスコア化し、その効果の認識と実際の使用頻度の差を求めることで、両者の関係を検討した。ハンガリー、台湾、韓国、そして中国という異なるL2学習環境で実施されたこれらの研究はいずれも、L2教員が効果的だと認識しているMSを必ずしも授業に取り入れているわけではないことを示している。L2教員によるMSの使用を妨げる要因として、クラスの人数の多さ(Cheng & Dörnyei, 2007; Guilloteaux, 2013),事務作業の多さ(Guilloteaux, 2013),学生に従順さを、教員に権威を求める文化的規範(Cheng & Dörnyei, 2007),および教員が学生の動機づけを優先事項としない姿勢(Guilloteaux, 2013)などが挙げられている。

こうした知見を踏まえ、Lee and Pun(2021)は、Dörnyei(2001)のMSを枠組みとして、香港の中等教育機関、大学や大学以外の高等教育機関に所属するL2教員58名への質問紙調査と20名へのグループインタビューを実施した。その結果、MSの使用に影響を与える要因には、学習者、教員、そして教育環境に起因する、計9つの要因があることを明らかにしている(表1)。

■表1: Lee and Pun (2021) で明らかになったL2教員のMS使用に影響を与える要因と回答数を参考に調査者が作成

| 起源   | 要因           | 要因の説明                                   | 回答数<br>(n= 313) |
|------|--------------|-----------------------------------------|-----------------|
| 学習者  | 学習者の受容(内容以外) | 学習内容と関連のない、MSによる活動の形式や雰囲気を学習者が受け入れるかどうか | 62              |
|      | 学習者の受容(内容)   | MSによって扱う、学習内容そのものを学習者が受け入れるかどうか         | 47              |
|      | 学習者のレディネス    | 学習者がMSの効果を楽しむ準備が整っているかどうか               | 40              |
|      | 学習者の行動       | MSの取り入れにより、学習者に望ましくない行動が生じないか           | 18              |
| 教員   | 教員のコスト       | 教員がMSを導入する際、どの程度コストがかかるか                | 46              |
|      | 教員の知識        | 教員がMS実施のために必要な知識を有しているか                 | 13              |
| 教育環境 | 授業内の実施       | 授業内でMSの実施が可能であるかどうか                     | 43              |
|      | 科目の要件        | MSと授業内容に親和性があるかどうか                      | 28              |
|      | 教育機関の要件      | MSと学校の方針に親和性があるかどうか                     | 16              |

Lee and Pun(2021)は様々な学校種のL2教員を対象とした調査を行ったが、教員のMS使用に影響を与える要因はコンテキストによって異なることが指摘されている(Ye & Hu, 2024)。このことを踏まえると、L2教員のMS使用に影響を与える要因を理解するには、特定のL2学習環境に焦点を当て、そこに身を置く教員を対象に調査を実施する必要があると考えられる。

### 3 本調査について

本調査の目的は、日本のEFL環境において、(調査1)英語教員のMSに関する効果の認識と彼らの実際のMS使用はどのような関係があるか、および(調査2)彼らのMS使用に影響を与える要因にはどのようなものがあるかを明らかにすることである。調査は申請者の所属する工業高等専門学校の英語教員を対象に実施し、(調査1)は量的質問紙による調査で、(調査2)は自由回答式の質問紙による質的調査で明らかにする。それぞれの調査におけるリサーチクエスチョン(RQ)は以下の通りである。

調査1:

RQ1 英語教員は、MSの効果についてどのように認識しているか。

RQ2 英語教員は、MSを授業においてどのくらいの頻度で使用しているか。

RQ3 MSの効果についての英語教員の認識と、彼らの授業におけるMS使用頻度の関係はどのようなものか。

調査2:

RQ4 英語教員のMS使用に影響を与える要因にはどのようなものがあるか。

### 4 調査1

#### 4.1 参加者

調査者の所属する近畿の工業高等専門学校で英語の授業を担当する教員6名(常勤4名, 非常勤2名, 計6名)を対象に本調査を実施した。高等専門学校(以下, 高専)とは、実践的・創造的技術者の養成を目的とした5年制の高等教育機関である(文部科学省, n.d.)。当該高専における英語授業は、1年生から3年生までは文部科学省指定の英語コミュニケーションや論理表現の検定教科書を用いて行われ、4, 5年生の授業では工学関連の英語教材やTOEIC対策用の大学生向けテキストが使用されている。6名の教員は英語を母国語としない日本人教員で、本調査実施時点では、当該高専における教授経験が4年以上あり、高専生や高専という学習環境について熟知していた。

#### 4.2 調査項目

調査には、以下2種類の質問項目を用いた。

##### (A)MSの効果の認識について尋ねる項目

Sugita(2009)の65項目から成るMSのリストを用いて、各MSが学生の英語学習への動機づけにどの

程度効果的だと思うかを教員に尋ねた(資料1)。回答は、6:「とても効果的だと思う」から 1:「まったく効果的だと思わない」までの6段階のリッカート・スケールで評価された。偶数段階の評価を採用した理由は、5段階評価などの奇数段階を用いると、日本人の回答者は中央値(例:5段階評価では3)を選択しやすく、その結果分散が小さくなる傾向があるためである(Sugita-McEown & Takeuchi, 2014)。本尺度の信頼性を示すクロンバックのアルファ係数は .96だった。

#### (B)MSの授業における使用頻度について尋ねる項目

(A)同様、Sugita(2009)のリストを用いて、各MSを授業でどのくらいの頻度で取り入れているかを教員に尋ねた(資料2)。回答は、6:「頻繁に取り入れている」から 1:「まったく取り入れていない」までの6段階のリッカート・スケールで評価した。本尺度のクロンバックのアルファ係数は、.95だった。

### 4.3 実施手順

データ収集に先立ち、調査者の所属する高専から調査実施の許可を得た。各教員に調査者が調査の内容について説明し、調査への参加は任意であることや、得られたデータの守秘義務について周知した。さらに、本調査は、Cheng and Dörnyei(2007)が指摘する「社会的望ましきバイアス」を軽減する措置を講じて実施した。社会的望ましきバイアスとは、回答者が実際よりも好ましい行動を報告する傾向を指し、「彼らがポジティブまたは期待される回答だと信じる内容に沿って回答する」ことを意味する(Cheng & Dörnyei, 2007, p. 156)。本調査では、(A)と(B)の調査を同時に行うと、教員が実際の使用頻度を過大または過小に報告する可能性が考えられた。これを防ぐため、(A)と(B)の調査は別の日に実施し、教員が過去の回答を参照できないよう配慮した。

### 4.4 結果

#### 4.4.1. RQ1. 英語教員は、MSの効果についてどのように認識しているか

調査に参加した英語教員から得られた(A)についての回答データを、Kawamitsu and Takeuchi(2024)の5因子に基づいて分類し、各因子の平均値を算出した(表2)。その結果、英語教員が効果的だと認識しているMSは平均値の高い順に、「外発的動機づけの促進」、「達成感の得られやすい授業展開」、「興味・関心の刺激」、「自律のサポート」そして「学生主体の活動」だった。

■表2: MSの効果に関する英語教員の認識の平均(降順)

| 順位 | 因子名            | 平均(標準偏差)    |
|----|----------------|-------------|
| 1  | 外発的動機づけの促進     | 4.75 (0.76) |
| 2  | 達成感を得られやすい授業展開 | 4.54 (0.72) |
| 3  | 興味・関心の刺激       | 4.43 (0.35) |
| 4  | 自律のサポート        | 4.19 (0.79) |
| 5  | 学生主体の活動        | 4.14 (0.43) |

#### 4.4.2. RQ2. 英語教員は、MSを授業においてどのくらいの頻度で使用しているか

(A)と同様に、英語教員から得られた(B)の回答データを、Kawamitsu and Takeuchi(2024)の5因子に基づいて分類し、各因子の平均値を算出した(表3)。平均値を降順で並べた結果、MSの効果の認識についての結果(表2)と、3位と4位の因子が入れ替わっている点を除き、順位は一致していた。

■表3: 英語教員のMS使用頻度の平均(降順)

| 順位 | 因子名            | 平均(標準偏差)    |
|----|----------------|-------------|
| 1  | 外発的動機づけの促進     | 4.67 (0.88) |
| 2  | 達成感を得られやすい授業展開 | 4.11 (0.93) |
| 3  | 自律のサポート        | 3.91 (0.88) |
| 4  | 興味・関心の刺激       | 3.42 (0.41) |
| 5  | 学生主体の活動        | 2.99 (0.88) |

#### 4.4.3. RQ3. MSの効果についての英語教員の認識と、彼らの授業におけるMS使用頻度の関係はどのようなものか

表2と表3の平均値を用いて、MSの効果に関する教員の認識とMS使用頻度の関係を検討した。サンプルサイズが小さいため、Hiver et al.(2020)を参考に、スピアマンの順位相関分析とベイズのピアソン相関分析を併用した。帰無仮説の棄却に関しては、ベイズ統計を用いた分析の方が推測統計よりも厳密であるとされる(Norouzian et al., 2019)ことから、両分析において相関が支持された場合にのみ「相関がある」と判断した。なお、ベイズ相関分析では、竹内・水本(2023)を参考に、 $BF_{10} > 3$ を相関ありと判断する指標とした。

分析の結果、効果の認識と使用頻度間には、順位相関分析では相関が見られた( $r_s = .90$ ,  $p = .037$ ) (表4)が、ベイズ相関分析では相関関係が支持されなかった( $r = .80$ ,  $BF_{10} = 1.60$ ) (表5)。また、MSの5因子それぞれについて効果の認識と使用頻度の関係を検討したが、スピアマンの順位相関分析とベイズのピアソン相関分析の両方で相関が支持された因子は確認されなかった(表6から表10)。

■表4: MSの効果に関する英語教員の認識と彼らのMS使用頻度の関係(スピアマンの順位相関)

|          | MS使用頻度 |
|----------|--------|
| MSの効果の認識 | .90*   |

Note. \*  $p < .05$

■表5: MSの効果に関する英語教員の認識と彼らのMS使用頻度の関係(ベイズのピアソン相関)

|          |               | MS使用頻度 |
|----------|---------------|--------|
| MSの効果の認識 | Pearson's $r$ | .80    |
|          | $BF_{10}$     | 1.60   |

■表6: MS因子「学生主体の活動」の効果に関する英語教員の認識と使用頻度の関係

|       |            |                  | 使用頻度 |
|-------|------------|------------------|------|
| 効果の認識 | スピアマンの順位相関 | Spearman's $r_s$ | .11  |
|       |            | $p$              | 0.83 |
|       | ベイズのピアソン相関 | Pearson's $r$    | .18  |
|       |            | $BF_{10}$        | 0.52 |

■表7: MS因子「自律のサポート」の効果に関する英語教員の認識と使用頻度の関係

|       |            |                 | 使用頻度 |
|-------|------------|-----------------|------|
| 効果の認識 | スピアマンの順位相関 | Spearman's $rs$ | .77  |
|       |            | $p$             | 0.10 |
|       | ベイズのピアソン相関 | Pearson's $r$   | .75  |
|       |            | $BF_{10}$       | 1.70 |

■表8: MS因子「達成感を得られやすい授業展開」の効果に関する英語教員の認識と使用頻度の関係

|       |            |                 | 使用頻度 |
|-------|------------|-----------------|------|
| 効果の認識 | スピアマンの順位相関 | Spearman's $rs$ | .79  |
|       |            | $p$             | 0.06 |
|       | ベイズのピアソン相関 | Pearson's $r$   | .63  |
|       |            | $BF_{10}$       | 1.05 |

■表9: MS因子「興味・関心の刺激」の効果に関する英語教員の認識と使用頻度の関係

|       |            |                 | 使用頻度 |
|-------|------------|-----------------|------|
| 効果の認識 | スピアマンの順位相関 | Spearman's $rs$ | .15  |
|       |            | $p$             | 0.77 |
|       | ベイズのピアソン相関 | Pearson's $r$   | .05  |
|       |            | $BF_{10}$       | 0.49 |

■表10: MS因子「外発的動機づけの促進」の効果に関する英語教員の認識と使用頻度の関係

|       |            |                 | 使用頻度 |
|-------|------------|-----------------|------|
| 効果の認識 | スピアマンの順位相関 | Spearman's $rs$ | .49  |
|       |            | $p$             | 0.32 |
|       | ベイズのピアソン相関 | Pearson's $r$   | .60  |
|       |            | $BF_{10}$       | 0.98 |

#### 4.5 調査1の考察

MSの効果の認識と使用頻度の相関関係は、MS全体としても因子別としても支持されなかった。本調査で用いた分析手法は先行研究とは異なるものの、他のL2学習環境で行われた調査(Cheng & Dörnyei, 2007; Dörnyei & Csizér, 1998; Guilloteaux, 2013; Ye & Hu, 2024)と同様、本調査の対象とした高専においても、MSの効果を高く認識していても、そのMSが十分に使用されているとは限らないことが示された。

## 5 調査2

### 5.1 参加者

調査1と同じ参加者を対象に本調査を実施した。

### 5.2 調査項目

英語教員のMS使用に影響を与える要因を明らかにするために、Lee and Pun(2021)を参考に、以下の自由回答式の質問項目を用いて、参加者にそれぞれの項目についてできるだけ多く回答してもらえるよう依頼した。

- (1)-1 学生が英語の授業でやる気を出すために、授業でどんな工夫をしていますか。
- (1)-2 その方法を授業に取り入れることができている理由や背景を、できるだけ詳しく教えてください。
- (2)-1 学生が英語の授業でやる気を出すために使おうと考えたけれど、結局使わなかった(使わなくなった)方法を教えてください。
- (2)-2 その方法を授業に取り入れるのをやめた理由や背景について、できるだけ詳しく教えてください。

### 5.3 実施手順

データ収集に先立ち、調査者の所属する高専から調査実施の許可を得た。各教員に調査者が調査の内容について説明し、調査への参加は任意であることや、得られたデータの守秘義務について周知した。調査は、質問項目を記載したMicrosoft Word ファイルを参加者にメールで送付し、回答済みファイルを返信してもらう方法で実施した。なお、回答内容に不明瞭な点があった場合には、該当する教員に後日口頭で補足説明を依頼し、情報を補完した。

### 5.4 予備調査

5.2の(1)-1、(2)-1について教員から得られた回答、すなわち授業での使用が検討されたMS( $n=21$ )について、それらがDörnyei(2001)のマクロストラテジーに含まれるかを検討した。この分析は、L2教員のMS使用に影響を与える要因を調査したLee and Pun(2021)がDörnyei(2001)のMSリストをフレームワークにしており、本調査との比較を可能にするために行ったものである。分析の結果、回答されたすべてのMSがDörnyei(2001)のフレームワークに収まることが確認された(表11)。

■表11: Dörnyei (2001) のマクロストラテジーと本調査で回答されたMS との対応関係

| Dörnyei (2001) のマクロストラテジー                                   | 回答数<br>( $n=21$ ) |
|-------------------------------------------------------------|-------------------|
| 1. 扱う教材に対する自分の熱意と、それが自分に個人的にどんな影響を及ぼしているかについて、実例を挙げて説明し解説する | 3                 |
| 6. 集団の結束力強化を促進する                                            | 1                 |
| 11. L2とその使用者、また外国らしさ全般に対する肯定的で開放的な気質を育てることで、「統合的」価値観を高める    | 4                 |
| 13. 特定の課題および学習全般に関する生徒の成功期待感を高める                            | 9                 |
| 15. 教育課程と教材を、学習者に関連の深いものにする                                 | 1                 |
| 17. 教室での活動の単調さを打破することによって、学習をより興味深く楽しいものにする                 | 1                 |
| 18. タスクの魅力を増すことにより、学習を学習者にとって興味深く楽しいものにする                   | 1                 |
| 32. 肯定的情報フィードバックを学習者に与える                                    | 1                 |

## 5.5 結果

5.2の(1)-2,(2)-2について教員から得られた回答を, Lee and Pun(2021)を参考にコード化し, 同研究で明らかになったMS使用に影響を与える要因に該当するものと, 該当しないものに分類した。得られたコードは計33であり, そのうち27はLee and Pun(2021)の分類に該当した(表12), 一方, 残り6つのコードは, 起源としてはLee and Pun(2021)の分類に含まれるものの, 要因としては分類できなかった(表13)。

■表12: Lee and Pun (2021) の要因の分類と, 本調査で回答されたMS, MS使用の理由や背景に関する回答やコードの対応関係

| Lee and Pun(2021)の分類 |                  |                |                                                      |                |
|----------------------|------------------|----------------|------------------------------------------------------|----------------|
| 起源                   | 要因               | MS・理由や背景・コードの例 |                                                      | 回答数<br>(n= 27) |
| 学習者                  | 学習者の受容<br>(内容以外) | MS             | 授業の最初にテーマを設定してペアで英会話する時間を設定する                        | 7              |
|                      |                  | 理由・背景          | 授業初めの緊張を解きほぐすため                                      |                |
|                      |                  | コード            | リラックスできる雰囲気づくり                                       |                |
|                      | 学習者の受容<br>(内容)   | MS             | 本時で扱う内容ははじめに説明する                                     | 10             |
|                      |                  | 理由・背景          | 学生にとって本時で取り扱う内容を最初に知ることが授業を受ける際に聞くポイントの緩急をつけられると思うから |                |
|                      |                  | コード            | 学習ポイント把握の助け                                          |                |
|                      | 学習者の<br>レディネス    | MS             | 文法解説の動画で予習させ, 授業内ではコミュニケーション活動に集中させる                 | 1              |
|                      |                  | 理由・背景          | ターゲットとしている学習者の学力が比較的高いため, 単調な解説は授業外に持ち込むことが可能なため     |                |
|                      |                  | コード            | 学力の高さ                                                |                |
|                      | 学習者の行動           | MS             | 該当なし                                                 | 0              |
|                      |                  | 理由・背景          | 該当なし                                                 |                |
|                      |                  | コード            | 該当なし                                                 |                |
| 教員                   | 教員のコスト           | MS             | パワーポイントのスライドに英文読解に必要な情報をすべて載せる(取りやめ)                 | 6              |
|                      |                  | 理由・背景          | パワーポイント作成に多くの時間を割かなければならず, 直前まで準備が必要だから              |                |
|                      |                  | コード            | 準備時間の捻出の難しさ                                          |                |
|                      | 教員の知識            | MS             | 映画の視聴(取りやめ)                                          | 1              |
|                      |                  | 理由・背景          | 伝えたいことがありすぎて, 映画を見せるにあたり焦点を絞ることができなかった               |                |
|                      |                  | コード            | 学習の焦点を決めることの難しさ                                      |                |
| 教育<br>環境             | 授業内の実施           | MS             | 授業の間にクイズを入れる(取りやめ)                                   | 2              |
|                      |                  | 理由・背景          | 授業の進捗との関連で, 時間を捻出するのが難しい                             |                |
|                      |                  | コード            | 実施時間確保の難しさ                                           |                |
|                      | 科目の要件            | MS             | 該当なし                                                 | 0              |
|                      |                  | 理由・背景          | 該当なし                                                 |                |
|                      |                  | コード            | 該当なし                                                 |                |
|                      | 教育機関の<br>要件      | MS             | 該当なし                                                 | 0              |
|                      |                  | 理由・背景          | 該当なし                                                 |                |
|                      |                  | コード            | 該当なし                                                 |                |

■表13: Lee and Pun (2021) の分類に該当しなかったMS使用の理由や背景に関する6コード

| 回答されたMS・理由や背景 |       |                                           | コード                  | 起源       |
|---------------|-------|-------------------------------------------|----------------------|----------|
| 1             | MS    | 英語の歌を使ったリスニング練習(取りやめ)                     | 学習者へのサポートの<br>必要性の消失 | 学習者      |
|               | 理由・背景 | スマホなどを使って、学生自身で手軽にできるようになったから             |                      |          |
| 2             | MS    | 本時で扱う内容をはじめに説明する                          | 教員の進度判断の助け           | 教員       |
|               | 理由・背景 | 教員は本時のメニューを発表することで進度に目処が立つから              |                      |          |
| 3             | MS    | 授業プリントを冊子にして配る                            | 教員の教材管理の楽さ           | 教員       |
|               | 理由・背景 | 教員がプリントを整理・管理することが苦手で、冊子にしておく<br>と管理が楽だから |                      |          |
| 4             | MS    | 自作のパワーポイントスライドの活用                         | 教員の教材の蓄積             | 教員       |
|               | 理由・背景 | 蓄積ができるのがよいから                              |                      |          |
| 5             | MS    | 情報共有ツールを使用したリアルタイムの質問対応(取りやめ)             | 教員の授業スタイルの変化         | 教員       |
|               | 理由・背景 | 解説をすべて動画配信に切り替えたため                        |                      |          |
| 6             | MS    | オンライン教材の使用(取りやめ)                          | ネットワーク環境の悪さ          | 教育<br>環境 |
|               | 理由・背景 | 教室でインターネット接続が上手くいかなかった                    |                      |          |

## 5.6 調査2の考察

本調査で得られたコードのうち8割以上はLee and Pun(2021)の分類に該当したが、同研究における学習者要因の「学習者の行動」、および教育機関要因の「科目の要件」や「教育機関の要件」に関するコードは本調査においては確認されなかった。「学習者の行動」に関するコードが得られなかったのは、当該高専において学生の自律が重視されている文化的背景を反映しているものと考えられる。また、「科目の要件」や「教育機関の要件」に関するコードが確認されなかったのは、当該高専においては英語授業の内容について教員に大きな裁量を与えられているためであると推察される。

さらに、本調査で得られたコードのうち、Lee and Pun(2021)の分類に該当しなかった6コードは、当該高専特有の要因であると考えられる。表14は、表13のコードを参考に命名した高専特有の要因を示している。当該高専においては、Lee and Pun(2021)の分類に該当した要因のほか、MSの取り入れを検討する際に学生へのサポートの必要性や、MSを取り入れによる教員側のメリット、授業スタイル、学校設備の充実度といった要因も考慮していることが示唆された。表15は、表13および表14をもとに、本調査で対象とした高専において英語教員のMS使用に影響を与えられと考えられる要因をまとめたものである。

■表14: 教員のMS使用に影響する高専特有と考えられる要因

| 起源   | コード                                  | 要因名           |
|------|--------------------------------------|---------------|
| 学生   | 学習者へのサポートの必要性の消失                     | 学習者へのサポートの必要性 |
| 学生   | 教員の進捗判断の助け<br>教員の教材管理の楽さ<br>教員の教材の蓄積 | 教員のメリット       |
|      | 教員の授業スタイルの変化                         | 教員の授業スタイル     |
| 教育機関 | ネットワーク環境の悪さ                          | 学校設備          |

■表15: 調査の対象とした高専において英語教員のMS使用に影響を与える要因のまとめ

| 起源   | 要因                                                       |
|------|----------------------------------------------------------|
| 学習者  | 学習者の受容(内容以外)<br>学習者の受容(内容)<br>学習者のレディネス<br>学習者へのサポートの必要性 |
| 教員   | 教員のコスト<br>教員の知識<br>教員のメリット<br>教員の授業スタイル                  |
| 教育機関 | 授業内の実施<br>学校設備                                           |

## 6

## 調査のまとめ

本調査では、調査者の所属する高専において、(調査1)英語教員によるMSの効果の認識とMS使用頻度の関係、および(調査2)彼らのMS使用に影響を与える要因を検討した。調査1の結果、先行研究(Cheng & Dörnyei, 2007; Dörnyei & Csizér, 1998; Guilloteaux, 2013; Ye & Hu, 2024)と同様に、教員によるMSの効果の認識とMS使用頻度の関係は必ずしも直線的なものではないことが明らかになった。調査2では、Lee and Pun(2021)の分類に該当する要因が多く見られた一方で、同研究の一部の要因には対応する回答が見られず、また本調査が対象とした高専特有のものだと考えられる要因も確認された。

本調査の限界を述べる。調査2で明らかになった要因(表15参照)について、それぞれの変数が実際にどの程度MSの使用に影響を与えているのかは明らかになっていない。各要因とMS使用との関係を明らかにするためには、表15に基づいて変数を設定し、量的調査を行う必要がある。相関関係を検討するのであれば、調査1で採用したスピアマンの順位相関分析やペイズのピアソン相関分析が適切だろう。しかし、L2教員の授業実践が多様な要因の相互作用や対立の上に成り立っていることを踏まえると(Borg, 2003),

本調査で確認された要因同士が複雑に関わり合っているケースも想定される。例えば、相互作用の例としては、「MSを用いた活動が学生に合っており、かつ授業内で実施可能と判断された場合に使用される」ケースがある。対立の例としては、「学生に合っていると認識しているが、コストがかかるため使用しない」または「コストはかかるが、学生に合っているため使用する」といった状況が考えられる。このように複数の変数間の複雑な関係を分析するには、例えば構造方程式モデリング(SEM)の手法が有効であろう。ただし、この手法を採用するには十分なサンプルサイズが必要であり、本調査のように特定のL2学習環境に焦点を当てた調査では採用が難しいと考えられる。十分なサンプルサイズを確保するためには、比較的学习環境が似ている、例えば他の高専も対象に含めた調査を実施することが、現実的かつ有益なアプローチだと考えられる。

## 謝辞

本研究を発表する貴重な機会を与えてくださった公益財団法人 日本英語検定協会と関係者の皆様、選考委員の先生方に深く御礼申し上げます。特に助言担当の寺内 一先生には、今後の調査の方向性に関する貴重なアドバイスをいただき大変感謝しております。最後に、調査に協力してくださった大阪公立大学工業高等専門学校の英語教員の皆様に深く御礼申し上げます。なお、本報告書に関して、開示すべき利益相反関連事項はありません。

## 引用文献

- Agawa, T. & Takeuchi, O. (2017). Pedagogical intervention to enhance self-determined forms of L2 motivation: Applying self-determination theory in the Japanese university EFL context. *Language Education & Technology*, 54, 135-166.
- Alrabai, F. (2016). The effects of teachers' in-class motivational intervention on learners' EFL achievement. *Applied Linguistics*, 37(3), 307-333.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81-109.
- Cheng, H. F., & Dörnyei, Z. (2007). The use of motivational strategies in language instruction: The case of EFL teaching in Taiwan. *Innovation in Language Learning and Teaching*, 1(1), 153-174.
- Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge University Press.
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9-42). Multilingual Matters.
- Dörnyei, Z., & Csizér, K. (1998). Ten commandments for motivating language learners: Results of an empirical study. *Language Teaching Research*, 2(3), 203-229.
- Dörnyei, Z., Ibrahim, Z., & Muir, C. (2015). Directed motivational currents: Regulating complex dynamic systems through motivational surges. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 95-105). Multilingual Matters.
- Dörnyei, Z., & Ottó, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics*, 4, 43-69.
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. Routledge.
- Fukui, H. & Yashima, T. (2021). Exploring evolving motivation to learn two languages simultaneously in a study-abroad context. *The Modern Language Journal*, 105, 267-293.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Edward Arnold.
- Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second-language learning*. Newbury House.
- Guilloteaux, M. J. (2013). Motivational strategies for the language classroom: Perceptions of Korean secondary school English teachers. *System*, 41(1), 3-14.
- Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, 42(1), 55-77.
- Henry, A., Davydenko, S., & Dörnyei, Z. (2015). The anatomy of directed motivational currents: Exploring intense and enduring periods of L2 motivation. *Modern Language Journal*, 99(2), 329-345.
- 廣森友人 (2010). 「第3章 動機づけ研究の観点から見た効果的な英語指導法」小嶋英夫・尾関直子・廣森友人 (編)『成長する英語学習者 学習者要因と自律学習 (pp. 47-74)』. 大修館書店.
- 廣森友人・和田玲 (編著) (2024). 『エンゲージメントを促す英語授業—やる気と行動をつなぐ新しい動機づけ概念—』. 大修館書店.
- Hiver, P., Zhou, A., Tahmouresi, S., Sang, Y., & Papi, M. (2020). Why stories matter: Exploring learner engagement and

## 引用文献

- metacognition through narratives of the L2 learning experience. *System*, 91, 1-13.
- 河合靖 (2010). 「第2章 学習者の多様性-学習スタイル」小嶋英夫・尾関直子・廣森友人 (編) 『成長する英語学習者 学習者要因と自律学習 (pp. 21-46)』. 大修館書店.
- Kawamitsu, D. & Takeuchi, O. (2024). Students' and teachers' perceptions of L2 motivational strategies: A situated analysis in a Japanese EFL context. *Asian Journal of English Language Teaching*, 33, 9-28.
- Lee, T. S. O., Gardner, D., & Lau, K. (2020). The effects of L2 motivational strategies: within and beyond the L2 classroom. *Innovation in Language Learning and Teaching*, 14(5), 451-465.
- Lee, T. S. O., & Pun, D. H. Y. (2021). The feasibility of motivational strategies in language classrooms: A tentative teacher-oriented definition. *Teaching and Teacher Education*, 106, 1-11.
- Mercer, S., & Dörnyei, Z. (2020). *Engaging Language Learners in Contemporary Classrooms*. Cambridge University Press.
- 文部科学省 (2015). 『平成26年度英語力調査 (高校3年生) 結果の概要』 [https://www.mext.go.jp/component/a\\_menu/education/detail/\\_icsFiles/afidfile/2015/07/03/1358071\\_01.pdf](https://www.mext.go.jp/component/a_menu/education/detail/_icsFiles/afidfile/2015/07/03/1358071_01.pdf)
- 文部科学省 (2017). 『平成29年度英語力調査 (高校3年生) 結果の概要』 [https://www.mext.go.jp/a\\_menu/kokusai/gaikokugo/\\_icsFiles/afidfile/2018/04/06/1403470\\_03\\_1.pdf](https://www.mext.go.jp/a_menu/kokusai/gaikokugo/_icsFiles/afidfile/2018/04/06/1403470_03_1.pdf)
- 文部科学省 (n.d.). 『高等専門学校 (高専) について』 [https://www.mext.go.jp/a\\_menu/koutou/kousen/index.htm](https://www.mext.go.jp/a_menu/koutou/kousen/index.htm)
- Moskovsky, C., Alrabai, F., Paolini, S., & Ratcheva, S. (2013). The effects of teachers' motivational strategies on learners' motivation: A controlled investigation of second language acquisition. *Language Learning*, 63(1), 34-62.
- Noels, K.A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Orientations and self-determination theory. *Language Learning*, 50, 57-85.
- Norouziyan, R., Miranda, M. De, & Plonsky, L. (2019). A Bayesian approach to measuring evidence in L2 research: An empirical investigation. *Modern Language Journal*, 103(1), 248-261.
- 尾関直子 (2010). 「第4章 学習ストラテジーとメタ認知」小嶋英夫・尾関直子・廣森友人 (編) 『成長する英語学習者 学習者要因と自律学習 (pp. 75-103)』. 大修館書店.
- Papi, M., & Abdollahzadeh, E. (2012). Teacher motivational practice, student motivation, and possible L2 selves: An examination in the Iranian EFL context. *Language Learning*, 62(2), 571-594.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54-67.
- Sugita, M. (2009). *"I don't like English." A comprehensive study of motivating EFL students in the Japanese secondary school context*. Unpublished doctoral dissertation. Kansai University.
- Sugita, M., & Takeuchi, O. (2010). What can teachers do to motivate their students? A classroom research on motivational strategy use in the Japanese EFL context. *Innovation in Language Learning and Teaching*, 4(1), 21-35.
- Sugita-McEown, M., & Takeuchi, O. (2014). Motivational strategies in EFL classrooms: How do teachers impact students' motivation? *Innovation in Language Learning and Teaching*, 8(1), 20-38.
- Taguchi, T., Magid, M., & Papi, M. (2009). The L2 motivational self system among Japanese, Chinese, and Iranian learners of English: A comparative study. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity, and the L2 self* (pp. 66-97). Multilingual Matters.
- 竹内理 (2010). 「学習者の研究からわかること一個別から統合へ」小嶋英夫・尾関直子・廣森友人 (編) 『成長する英語学習者 学習者要因と自律学習 (pp. 3-20)』. 大修館書店.
- 竹内理. (2016). 英語学習の動機を高め、維持するには —動機づけ要因と動機づけ方略の観点から—. *JACET Kansai Journal*, 18, 36-51.
- 竹内理. (2023). そこに動機はあるのか? —外国語学習と動機づけの関係—. 『通訳翻訳研究』23, 1-13.
- 竹内理・水本篤 (編) (2023). 『外国語教育研究ハンドブック—研究手法のより良い理解のために—【増補版】』. 松柏社.
- Ueki, M., & Takeuchi, O. (2013). Exploring the concept of the ideal L2 self in an Asian EFL context: The case of Japanese university students. *Innovation in Language Learning and Teaching*, 7, 238-252.
- Wong, R. M. H. (2013). An investigation of strategies for student motivation in the Chinese EFL context. *Innovation in Language Learning and Teaching*, 8(2), 132-154. <https://doi.org/10.1080/17501229.2013.777449>
- Yashima, T., Noels, K., Shizuka, T., Takeuchi, O., Yamane, S., & Yoshizawa, K. (2009). The interplay of classroom anxiety, intrinsic motivation, and gender in the Japanese EFL context. *Kansai University Journal of Foreign Language Education and Research*, 17, 41-64.
- Ye, X., & Hu, G. (2024). Teachers' stated beliefs and practices regarding L2 motivational strategies: A mixed-methods study of misalignment and contributing factors. *System*, 121, 1-22.

# 資料1：MSの効果の認識について尋ねる項目 .....

以下の 65 項目のそれぞれに関して、先生のご担当のクラスを頭に思い浮かべながら、  
下記の選択肢の中からあてはまる数字を選び、回答してください。

## 【選択肢】

1. 学生のやる気を高める上でまったく効果的だとは思わない
2. 学生のやる気を高める上で効果的ではないと思う
3. 学生のやる気を高める上でどちらかといえば効果的ではないと思う
4. 学生のやる気を高める上でどちらかといえば効果的だと思う
5. 学生のやる気を高める上で効果的だと思う
6. 学生のやる気を高める上でとても効果的だと思う

|                                                         |   |   |   |   |   |   |
|---------------------------------------------------------|---|---|---|---|---|---|
| 1. 英語の勉強法や外国に関する興味深い話をする                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. 英語学習から満足感が得られることを伝える                                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. 学生の英語学習の進歩を常に気にかけていることを伝える                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. わからないことがあればいつでも質問にくるように勧める                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. 学生が少しチャレンジできるくらいの一段高い目標を設定する                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. 英語を学習する上では間違えることも大切だと伝える                             | 1 | 2 | 3 | 4 | 5 | 6 |
| 7. 授業中にユーモアを取り入れる                                       | 1 | 2 | 3 | 4 | 5 | 6 |
| 8. 英語の勉強に適した教室の雰囲気を作る                                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 9. 授業のはじめに、学生の緊張を解くための楽しい活動をする                          | 1 | 2 | 3 | 4 | 5 | 6 |
| 10. 授業中に小グループを使って、いろいろな学生と交流する機会を与える                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 11. 教科書以外の教材を使う                                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 12. 学生にフィードバックを与える                                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 13. 英語が得意な学生と、そうでない学生をいっしょに活動させる                        | 1 | 2 | 3 | 4 | 5 | 6 |
| 14. 学生が楽しめそうな活動（歌やゲームなど）を中心に進める                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 15. 授業で文化的な要素（外国の生活について話すなど）を取り入れる                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 16. 英語学習に成功した有名人の体験談を話す                                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 17. 英語ができるようになったら、将来的に有利（良い仕事を得られるなど）<br>であることを学生に意識させる | 1 | 2 | 3 | 4 | 5 | 6 |
| 18. 英語が世界でどれだけ必要とされているかを伝える                             | 1 | 2 | 3 | 4 | 5 | 6 |
| 19. 学生に、実際の英語能力（中間、期末試験、TOEIC Bridge など）を意識させる          | 1 | 2 | 3 | 4 | 5 | 6 |
| 20. 授業中に予習や復習がきちんとできているか確認する                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 21. 学生が誰か（保護者など）に助けてもらえる環境にいるのかを確認する                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 22. その活動することによって英語力がどのように身につくのかを説明する                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 23. 学生が達成できそうな授業目標を定める                                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 24. 学生の望みや興味などをできるだけ授業に取り入れる                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 25. 日常生活（クラブ活動など）についての英語表現を教える                          | 1 | 2 | 3 | 4 | 5 | 6 |
| 26. 学生の英語学習に対する間違った考え（訳ができればよいなど）を積極的に直す                | 1 | 2 | 3 | 4 | 5 | 6 |
| 27. いろいろな英語学習法を学生に意識させて教える                              | 1 | 2 | 3 | 4 | 5 | 6 |

## (資料1 続き)

|                                                         |   |   |   |   |   |   |
|---------------------------------------------------------|---|---|---|---|---|---|
| 28. 教員自身が教授法や活動に対する考え方を多様化する                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 29. 学生のやる気を上げることに気を配る                                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 30. 学生が普段しない活動をときどき取り入れる                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 31. 作品として残る活動（日記や詩など）をする                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 32. それぞれの学生が頭もからだも使える活動をする                              | 1 | 2 | 3 | 4 | 5 | 6 |
| 33. 授業で学生にそれぞれ役割（発表や教員の補助当番など）を分担させる                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 34. 目標を自分で選択する余地を学生に与える                                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 35. 目標をいつまでに達成できればよいのか、締め切りを決める                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 36. 学生の伸びを見て、目標達成ができているかどうかをチェックする                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 37. 授業で学生が達成感を感じられる機会を増やす                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 38. 学生の能力に合ったレベルの活動を選択する                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 39. テストで、学生のやれないことよりやれることに焦点を置く                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 40. テストに発展問題を付け加える                                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 41. 英語ができるようになるためには、努力をすることが大切だと伝える                     | 1 | 2 | 3 | 4 | 5 | 6 |
| 42. ほかの学生と比べず、ひとりひとりの達成度を評価する                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 43. 学生同士で競争して勉強するように促す                                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 44. 学生同士で協力して勉強するように促す                                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 45. 学生に成績の評価基準をわかりやすく説明する                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 46. 英語がうまく話せないときのために、コミュニケーションの方法<br>（言い換えやつなぎ言葉など）を教える | 1 | 2 | 3 | 4 | 5 | 6 |
| 47. みんなの前ではめてあげられるようなチャンスを学生に与える                        | 1 | 2 | 3 | 4 | 5 | 6 |
| 48. 学生を批判することを避ける                                       | 1 | 2 | 3 | 4 | 5 | 6 |
| 49. 学生が予期しない形（いきなり指名するなど）で注目を浴びることを避ける                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 50. 同じ目標に向かって、グループがいっしょになってできる活動を設定する                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 51. 評価には個人の結果だけを入れるのではなく、グループでの結果も入れる                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 52. グループ内でよくやっていくために、グループ活動のルールや方法を教える                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 53. 学生のリーダーシップを尊重し、学生中心の授業を展開する                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 54. 学習意欲を上げることがいかに大切かを教える                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 55. 教員が役立つと思ったやる気を出す方法を学生に教える                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 56. 学生自身で学習意欲を上げる方法を考えるように促す                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 57. 失敗は能力が足りていないからではなく、努力が足りなかったからだと励ます                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 58. 失敗は能力が足りていないからではなく、間違った勉強法をしていたからだと励ます              | 1 | 2 | 3 | 4 | 5 | 6 |
| 59. どこに焦点を当てて学習すればよいかを教える                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 60. 学習の伸びを記録（グラフなど）して、学生の進歩を目に見える形でほめる                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 61. 学生が褒美だけにとらわれないように気をつける                              | 1 | 2 | 3 | 4 | 5 | 6 |
| 62. 何らかの目に見えるもの（シールなど）で褒美を与える                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 63. 成績には、学生が達成したことだけではなく、努力の過程も取り入れる                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 64. 筆記テスト以外（音読テストやインタビューテストなど）の評価も成績に入れる                | 1 | 2 | 3 | 4 | 5 | 6 |
| 65. 自己評価カードなどを使って、学生自身に自分の活動を評価させる                      | 1 | 2 | 3 | 4 | 5 | 6 |

資料2: MSの授業における使用頻度について尋ねる項目.....

以下の65項目のそれぞれに関して、先生のご担当のクラスを頭に思い浮かべながら、  
下記の選択肢の中からあてはまる数字を選び、回答してください。

【選択肢】

1. 自分の授業でまったく使用しない
2. 自分の授業でほとんど使用しない
3. 自分の授業であまり使用しない
4. 自分の授業でときどき使用する
5. 自分の授業で使用する事が多い
6. 自分の授業で頻繁に使用する

|                                                         |   |   |   |   |   |   |
|---------------------------------------------------------|---|---|---|---|---|---|
| 1. 英語の勉強法や外国に関する興味深い話をする                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. 英語学習から満足感が得られることを伝える                                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. 学生の英語学習の進歩を常に気にかけていることを伝える                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. わからないことがあればいつでも質問にくるように勧める                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. 学生が少しチャレンジできるくらいの一段高い目標を設定する                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. 英語を学習する上では間違えることも大切だと伝える                             | 1 | 2 | 3 | 4 | 5 | 6 |
| 7. 授業中にユーモアを取り入れる                                       | 1 | 2 | 3 | 4 | 5 | 6 |
| 8. 英語の勉強に適した教室の雰囲気を作る                                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 9. 授業のはじめに、学生の緊張を解くための楽しい活動をする                          | 1 | 2 | 3 | 4 | 5 | 6 |
| 10. 授業中に小グループを使って、いろいろな学生と交流する機会を与える                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 11. 教科書以外の教材を使う                                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 12. 学生にフィードバックを与える                                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 13. 英語が得意な学生と、そうでない学生をいっしょに活動させる                        | 1 | 2 | 3 | 4 | 5 | 6 |
| 14. 学生が楽しめそうな活動（歌やゲームなど）を中心に進める                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 15. 授業で文化的な要素（外国の生活について話すなど）を取り入れる                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 16. 英語学習に成功した有名人の体験談を話す                                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 17. 英語ができるようになったら、将来的に有利（良い仕事を得られるなど）<br>であることを学生に意識させる | 1 | 2 | 3 | 4 | 5 | 6 |
| 18. 英語が世界でどれだけ必要とされているかを伝える                             | 1 | 2 | 3 | 4 | 5 | 6 |
| 19. 学生に、実際の英語能力（中間、期末試験、TOEIC Bridge など）を意識させる          | 1 | 2 | 3 | 4 | 5 | 6 |
| 20. 授業中に予習や復習がきちんとできているか確認する                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 21. 学生が誰か（保護者など）に助けてもらえる環境にいるのかを確認する                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 22. その活動することによって英語力がどのように身につくのかを説明する                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 23. 学生が達成できそうな授業目標を定める                                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 24. 学生の望みや興味などをできるだけ授業に取り入れる                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 25. 日常生活（クラブ活動など）についての英語表現を教える                          | 1 | 2 | 3 | 4 | 5 | 6 |
| 26. 学生の英語学習に対する間違った考え（訳ができればよいなど）を積極的に直す                | 1 | 2 | 3 | 4 | 5 | 6 |
| 27. いろいろな英語学習法を学生に意識させて教える                              | 1 | 2 | 3 | 4 | 5 | 6 |

## (資料2 続き)

|                                                         |   |   |   |   |   |   |
|---------------------------------------------------------|---|---|---|---|---|---|
| 28. 教員自身が教授法や活動に対する考え方を多様化する                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 29. 学生のやる気を上げることにも気を配る                                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 30. 学生が普段しない活動をときどき取り入れる                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 31. 作品として残る活動（日記や詩など）をする                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 32. それぞれの学生が頭もからだも使える活動をする                              | 1 | 2 | 3 | 4 | 5 | 6 |
| 33. 授業で学生にそれぞれ役割（発表や教員の補助当番など）を分担させる                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 34. 目標を自分で選択する余地を学生に与える                                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 35. 目標をいつまでに達成できればよいのか、締め切りを決める                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 36. 学生の伸びを見て、目標達成ができているかどうかをチェックする                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 37. 授業で学生が達成感を感じられる機会を増やす                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 38. 学生の能力に合ったレベルの活動を選択する                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 39. テストで、学生のやれないことよりやれることに焦点を置く                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 40. テストに発展問題を付け加える                                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 41. 英語ができるようになるためには、努力をすることが大切だと伝える                     | 1 | 2 | 3 | 4 | 5 | 6 |
| 42. ほかの学生と比べず、ひとりひとりの達成度を評価する                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 43. 学生同士で競争して勉強するように促す                                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 44. 学生同士で協力して勉強するように促す                                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 45. 学生に成績の評価基準をわかりやすく説明する                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 46. 英語がうまく話せないときのために、コミュニケーションの方法<br>（言い換えやつなぎ言葉など）を教える | 1 | 2 | 3 | 4 | 5 | 6 |
| 47. みんなの前でほめてあげられるようなチャンスに学生を与える                        | 1 | 2 | 3 | 4 | 5 | 6 |
| 48. 学生を批判することを避ける                                       | 1 | 2 | 3 | 4 | 5 | 6 |
| 49. 学生が予期しない形（いきなり指名するなど）で注目を浴びることを避ける                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 50. 同じ目標に向かって、グループがいっしょになってできる活動を設定する                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 51. 評価には個人の結果だけを入れるのではなく、グループでの結果も入れる                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 52. グループ内でよくやっていくために、グループ活動のルールや方法を教える                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 53. 学生のリーダーシップを尊重し、学生中心の授業を展開する                         | 1 | 2 | 3 | 4 | 5 | 6 |
| 54. 学習意欲を上げることがいかに大切かを教える                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 55. 教員が役立つと思ったやる気を出す方法を学生に教える                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 56. 学生自身で学習意欲を上げる方法を考えるように促す                            | 1 | 2 | 3 | 4 | 5 | 6 |
| 57. 失敗は能力が足りていないからではなく、努力が足りなかったからだと励ます                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 58. 失敗は能力が足りていないからではなく、間違った勉強法をしていたからだと励ます              | 1 | 2 | 3 | 4 | 5 | 6 |
| 59. どこに焦点を当てて学習すればよいかを教える                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 60. 学習の伸びを記録（グラフなど）して、学生の進歩を目に見える形でほめる                  | 1 | 2 | 3 | 4 | 5 | 6 |
| 61. 学生が褒美だけにとらわれないように気をつける                              | 1 | 2 | 3 | 4 | 5 | 6 |
| 62. 何らかの目に見えるもの（シールなど）で褒美を与える                           | 1 | 2 | 3 | 4 | 5 | 6 |
| 63. 成績には、学生が達成したことだけではなく、努力の過程も取り入れる                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 64. 筆記テスト以外（音読テストやインタビューテストなど）の評価も成績に入れる                | 1 | 2 | 3 | 4 | 5 | 6 |
| 65. 自己評価カードなどを使って、学生自身に自分の活動を評価させる                      | 1 | 2 | 3 | 4 | 5 | 6 |

## 今回の報告について

～ 日本英語検定協会 名誉顧問より ～

名誉顧問 長 勝彦

高等学校のライティング指導における生成AIの活用とその効果  
—機能適切さ(Functional Adequacy)に着目して—

本年度は、AIを英語教育のテーマとして取り上げた研究報告が複数見られた。なかでも、渡邊大志先生の研究は、生成AIを高校のライティング指導に導入し、その教育的効果を実証的に検証した点に特色がある。研究に当たって、実に数多くのAIに関する文献から引用して研究報告を作成している。今後、AIに関する英語教育研究・AIを取り入れた英語授業実践など、AIに関して調査する際、少なくとも、渡邊大志先生が本研究報告書にて引用した文献一覧は大変参考になる。

来年度よりの「英検」研究助成報告に、研究部門、実践部門、調査部門にて、AIの活用とその効果等に関する多くの報告書の応募を期待します。

最後に、日本全国の英語研究会にて、可能な限り渡邊先生の今回のテーマを取り入れた授業を公開して、英語教育に貢献して欲しい。



名誉顧問 村木 英治

ChatGPTについての研究

つい最近私もChatGPTをやり始めた。試しに自分の名前を入れ「この人？」と聞いてみた。ネット空間に挙がっている私についての文献はほとんど記憶しているのだろう。敬意を感じさせながらそれらをうまくまとめて表現してくれた。気をよくして、「あなたは英語義務教育に役立てますか？」と尋ねてみた。ChatGPTの反応はすこぶる早い。意外な例も含まれている。便利なアプリである。

しかしその応答には責任感が感じられない。無人間レベルで動く生成AIには、教師と生徒という関係が重要な教育現場に応用される方策の一つ一つは注意深く実行されるべきであるという責任感はないのかもしれない。今回は大和田氏の論文を含めて3篇も生成AIの研究があった。それだけ議論が切迫されているのだろう。それに答えられるのは、しかし教育問題に愛情をもって取り組める研究者だけだ。AI工学分野に確固とした力を持ち、かつ教育問題にも熱意をもって取り組める研究者がここから育つことを願っている。それがChatGPT自身の教育にも役に立つ。

---

## 名誉顧問 小池 生夫

### 研究助言の長い年月を経て思うのはすべて感謝である

私は最近老化現象が形を成すようになってきた。今年度は研究部門5, 実践部門5, 調査部門1件の研究論文を読むのに、本来なら不必要な長い時間がかかるようになった。また、論文執筆者全員が第一級の研究者であると感じた。

本年度の特徴は、すべての部門でさまざまな数理統計モデルを利用して説得力のある論文を書いていることだ。明解な分析をみせてくれた。この傾向は、初期は研究部門だけの特許販売であったが、回数を数えるうちにこの傾向がはっきりし、ついに3部門全体の研究に及ぶことになった。これは今後しばらく続くであろう。そして次の時代は、「生成AI」の利用に関する研究であろう。そしてさらに新しい多様なモデルを使つての研究発表が一般的になるであろう。この傾向は説得力があり、助言者の助言でも、それに積極性を持たせようという意図がみえる。研究者、助言者ともに、数式によってその特徴を明らかに引き出すことをよしとする研究分析方法が、全体的に歓迎されるようになる。そういう傾向が決定的になったことを意味する。ある特定の結果が数字、数式の利用によって、より明白になったことを示すものといえよう。初代の助言者、指導者のひとりとして、ついにこの段階まできたかと思う。今後この傾向は一層明白になり、何のモデルを利用するのが適切かを検討することが研究者本人ばかりでなく、指導助言者の興味を引くことになる。これが新しい傾向になる。

これがよいかどうかは慎重に考慮することとし、まずは納得することになろう。私は長年指導助言をしてきた者のひとりとして現役の研究者、指導助言者に対し、その努力を尊重し、激励し、深く感謝の気持ちを表したい。

応募研究者、各指導者、その他の助言者の皆さん、本当にご苦労さま。皆さん相当のエネルギーを注いでくださった。その初期から今日まで、その真摯なる才智と努力に対し高く評価したい。それは私自身よく認識してきたからである。最後に事務局、研究者、研究助言者諸氏、今後とも才智と誠実なる努力をこのプログラムに注がれるよう望んでやまない。







ISSN 1348-7949

**EIKEN BULLETIN** vol. 37 2025

非売品

2026年2月7日 初版 発行

編集・発行 公益財団法人 日本英語検定協会

〒162-8055 東京都新宿区横寺町 55

<https://www.eiken.or.jp>

編集・組版／株式会社 ITP

印刷・製本／日新印刷株式会社

無断複製・転載を禁じます。 ©2026 公益財団法人 日本英語検定協会