

EIKEN BULLETIN

vol. 36 2024

第36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36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
- 上田 精也 全国高等学校長協会 英語国際部会代表・国際教育交流委員会 委員長
- * 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
- 富高 雅代 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

教科としての英語か,世界で通用する英語か

—グローバル社会で使える英語力の獲得を目指して—

横浜国立大学 教授

齊田 智里

秋は、日本全国の学校等で研究大会・公開授業が目白押しだ。中でも英語科の公開授業には多くの参加者が集まる。学習指導要領に基づき、児童生徒の主体性を重視しながら、発信スキル育成型の授業が展開される。今話題のAIもうまく活用している。十分に練られた指導案に、活発に言語活動に取り組む児童生徒の様子から、教科としての英語は十分に存在感を高めている。しかし、英語科にはグローバルな風が吹きつける。ある民間会社の2024年「英語能力指数」世界ランキングによると、日本は過去最低の92位だった。受験者はたまたまその会社のオンライン英語試験を受けた人たちで、その国民を代表しているわけではない。しかし、日本人の英語力が低いといわれて英語教育の研究者として無視するわけにはいかない。同様のことは、ETSのTOEFL iBTのスコアサマリーでもいえる。日本人受験者の平均スコアは、近隣のアジア諸国と比べてもかなり低い。スピーキングとライティングは過去10年間低水準を維持したままである。受験者はその国民を代表しているわけではなく、国別の比較自体意味がないことは百も承知の上で、このままでよいのか、と思ってしまう。私が第15回「英検」研究助成報告(2003年)に「高校入学時の英語能力値の年次推移」を発表し、英語力低下を指摘してから20数年が経過した。日本人の英語力は向上しているのだろうか？

令和5年度「全国学力・学習状況調査」で英語の平均正答率は国語や数学よりも低く、特に話すことの正答率が12.4%と非常に低かったことが、衝撃を持って報道された。問題が難しすぎたことも要因だろう。一人一台端末によるオンラインテストに不慣れなこともあっただろう。採点基準が厳しすぎたということもあるだろう。しかしそれで済む話だろうか？ OECDの2025年度PISA調査に、初めて英語力の国際比較調査が導入される。日本はこの調査に参加しない。15歳児を対象とした読解力、数学的リテラシー、科学的リテラシーの3分野のPISA調査には第1回から参加し、その結果に一喜一憂しながらも教科力を高め、3分野全てにおいて日本は世界トップレベルの水準を達成している。冒頭に述べたように、教科としての英語は立派である。国語、数学、理科に比べて決して劣るような授業をしているわけではない。世界で通用する英語力を児童生徒に獲得させるためにはどうしたらよいのか？「英検」研究助成報告書をそのために活用してほしい。関心のある論文をみつけて、一つでもいいから、実践に応用してほしい。堂々と英語でやりとりができ、世界で自信をもって生きていける次世代の日本人をみんなで育てていきたいと強く思う。

Contents

● はじめに

横浜国立大学 教授 齊田 智里 …… 3

● 報告別講評 専門選考委員 …… 6

A 研究部門

英語学習者エンゲージメントに関する文脈的モデルの妥当性検証

研究者：檜村 祐志 東京都／明治大学大学院 在籍 …… 12

結束性指標に基づく日本人英語学習者のエッセイライティングスコアの予測

研究者：久保 佑輔 茨城県／筑波大学大学院 在籍 …… 27

英検の長文テキストの読解に接辞の知識はどれくらい必要か？

ー Morpholex Affix Profiler を用いた検討 ー

研究者：駒野 樹 茨城県／筑波大学大学院 在籍 …… 43

B 実践部門

ディベートスキルとジャッジスキルの向上を目指した高校生英語ディベート初心者への効果的なフィードバック

研究者：坂口 寛子 福岡県／福岡県立春日高等学校 教諭 …… 59

Contents

明示的知識を外化して正確性の向上を目指す指導の効果

— 長期的なランゲージング・エピソードの分析を中心に —

研究者：高木 哲也 東京都／筑波大学附属高等学校 教諭 …… 79

ディスカッション活動において高校生の発話の量と質を高める取り組み

研究者：久山 慎也 広島県／広島県立安芸府中高等学校 教諭 …… 99

用法基盤モデルに基づいたスピーキング指導が中学生の『即興力』育成に

及ぼす効果の検証

研究者：吉澤 孝幸 秋田県／秋田県立秋田南高等学校 教諭 …… 111
東京都／昭和女子大学大学院 在籍



調査部門

日本語母語英語学習者が使用する定型表現の分析

— 「話すこと [やりとり]」と「話すこと [発表]」の技能育成をめざして —

研究者：小出 凱渡 茨城県／筑西市立明野中学校 教諭 …… 138
茨城県／筑波大学大学院 在籍

機械翻訳(MT)を学習ツールとするための考察

— 英作文を学ぶ大学生を対象とした量的・質的調査 —

研究者：湯浅 麻里子 大阪府／関西大学大学院 在籍 …… 159

● 今回の報告について 日本英語検定協会 名誉顧問より …… 179

*各研究者の所属は入選当時のものです。

報告別講評

A. 研究部門・報告Ⅰ

竹内 理

英語学習者エンゲージメントに関する 文脈的モデルの妥当性検証

【研究者：榎村 祐志】

本研究は、日本人英語学習者の学習エンゲージメントのモデルを動機づけの自己決定理論(Self-Determination Theory: SDT)の枠組みの下、先行研究に十分にあたり、堅実かつ精緻な統計手法を用い、総計753名(中高生403名、大学生350名)もの参加者を得て検証した力作である。特筆すべき点としては、アウトカム(スキル別の英語習熟度)までを含んでモデル化したことと、学習文脈(中高生・大学生)を考慮に入れて分析(多母集団モデルによるSEM)を行ったことにある。また、心理ネットワークモデルを利用して、変数間(なかでも自律動機づけの3変数である内発的、統合的、同一視的)のコンステレーション、つまりポジティブに関連しあったシステムを明らかにしたことも高い評価に値すると考えられる。

一方で、エンゲージメントの構成がagentiveを含む5側面ではなく4側面であることや、中学生の参加者が少ないため中高を一括りにしていること、英語習熟度が成績型ではなく自己評価型であること、さらにはエンゲージメントの階層性にまで拡げて研究がデザインされていないことなどは、著者も一部認識しているように、今後の改善点となるかもしれない。しかし、いずれも基礎研究としての本研究の価値を毀損するものではなく、本報告は高いクオリティーの論文と判断し、榎村氏の努力を讃えたい。SEMや心理ネットワークモデルは集団の特性に依拠する傾向もあるので、今後引き続き同種の研究を多様な参加者に協力を求めて行っていく、より堅固な知見を得ていかれることを期待したい。

A. 研究部門・報告Ⅱ

寺内 一

結束性指標に基づく日本人英語学習者の エッセイライティングスコアの予測

【研究者：久保 佑輔】

本研究は、日本人英語学習者が作成したエッセイに関して、テキスト一貫性とエッセイの質の調査を実施した。本研究により、内容語および形容詞の重複使用という指導しやすい事項が、エッセイの質に影響を与えるということが明確になったことは評価できる。これらは決して難しい技能ではなく、ともすれば心がけ一つで改善できることであるため、エッセイを書き始めたばかりの初級の学習者から指導可能である。学習の初期からエッセイの質を高めるための習慣をつけることができれば、自信を持ってエッセイを執筆できる自律した学習者の育成にもつながるだろう。

一方、筆者が最後に本研究の限界点と今後の課題でも言及しているように、本研究にはこれから発展させていくべき課題もある。特に、今回は結束性指標に焦点を絞っているが、語彙の洗練性や複雑性、文法の正確性などといった英語のライティングにおける主要な因子との関係については、本研究との相関も含めてさらなる検証を行っていただきたい。

さらに、考察の説得力を高めるために、下記の2つの方法を提案する。1つ目は、重回帰分析において強制投入法を実施したり、因果関係の推測のために共分散構造分析を実施したりすることである。本研究では用いられなかったこれらの統計手法を用いてデータを分析することにより、本研究で使用した手法の妥当性を高めることができるだろう。2つ目は、分析に使用したエッセイを、結果に基づいて文章を書き換えた上で分析し直すことである。その結果、スコアが変化するのであれば、筆者の考察は説得力が高まるはずである。

A. 研究部門・報告Ⅲ

齊田 智里

英検の長文テキストの読解に
接辞の知識はどれくらい必要か？

— Morpholex Affix Profiler を用いた検討 —

【研究者：駒野 樹】

学習指導要領外国語科の改訂により、学習すべき語彙数は大幅に増加している。駒野氏は、効率的な語彙学習において有用な知識として「接辞」に着目をし、英検の長文読解テキストを「接辞」の観点から分析した。分析した英文テキストは、英検の3級から1級までの5つの級の長文で、しかも、2014年度から2023年度の10年間に実施された全150回分の長文を分析したというから、その徹底ぶりから研究者としての意志の強さと実行力の高さが感じられる。対象となった総語数は225,210語である。英検のテストを対象とした接辞の大規模調査研究として、本研究は第一級の価値がある。

駒野氏は、Morpholex Affix Profilerを用いて、各級の長文テキストから抽出された接辞のレベルを分析した。具体的には、テキスト中に含まれる派生語の割合、登場する接辞の種類や頻度、語彙カバー率への寄与度を詳細に分析している。その結果、興味深い知見が次々と明らかになった。受験級が上がるにつれて、見出し語の割合は減少しているのに対し、派生語と屈折語の割合は増加傾向にあること、特に準1級から1級に関しては、派生語の割合が特に大きくなり、形態論的な複雑性が増していること、派生接辞のレベルは受験級が上がるにつれてその割合は増加していることから、難易度の高い文章の読解の際には高いレベルの接辞の知識が重要であること、特定の派生接辞が幅広い受験級の英文に高頻度で登場していること、語彙カバー率95%及び98%到達に必要な派生接辞の種類を各級ごとに明らかにしていることなどである。こうした駒野氏の緻密な分析から明らかになった貴重な知見が、語彙指導や英文読解テスト開発などに活かされることを期待したい。

B. 実践部門・報告Ⅰ

齊田 智里

ディベートスキルとジャッジスキルの
向上を目指した高校生英語ディベート
初心者への効果的なフィードバック

【研究者：坂口 寛子】

学習指導要領の改訂により、高等学校外国語科で「論理・表現」という新たな科目が設置され、教室での英語ディベート活動がこれまで以上に注目されるようになってきた。高等学校で長年、英語ディベートの指導経験を有する坂口氏は、結果のフィードバックの在り方に工夫の余地があることに気づき、本実践研究を行った。

生徒がディベーターのみならず審査員も経験する教室ディベートにおいては、生徒のディベート・スキルの向上とともに、ジャッジ・スキルの向上も必要となってくる。結果伝達の際、審査員からディベーターに口頭でフィードバックが与えられることが一般的であるが、坂口氏は加えて、フローシート、及び、グラフィックオーガナイザーによるフィードバックの方法に着目をし、ディベート初心者的高校生に対して、英語ディベート活動を通して3通りの異なるフィードバックを与えた。どのフィードバックがより有効かを検証するために、生徒のディベート・スキルとジャッジ・スキルが事前と事後でどのように変化したか、教師によるジャッジ判定の思考過程を生徒にわかりやすく伝えることができたかを、質問紙調査及び自由記述から明らかにしようとした。

ディベートやジャッジのスキルをまとめて分析をしてもよかったかもしれないが、スキルの各項目の変化に着目をして詳細に分析を行ったところに、坂口氏の実践家としてのこだわりがみられる。本実践研究を通して、英語ディベートでの口頭による通常のフィードバック以外のフィードバック方法にも光をあて、生徒の思考力、判断力、表現力等を一層育成するために、授業での多様な英語ディベート活動の可能性が示唆されたことの意義は大きい。

B. 実践部門・報告Ⅱ

西垣 知佳子

明示的知識を外化して正確性の向上を目指す指導の効果

— 長期的なランゲージング・エピソードの分析を中心に —

【研究者：高木 哲也】

第二言語習得におけるランゲージング (languageing) では、生徒が言語の規則や使用に関する自分の思考プロセスを言語化して外在化し、言語の構造や意味に対する理解を深める。本研究では、高校1年生61名を対象に30語程度のライティング指導を実施した。生徒の書いた英作文に対して、教師はエラー箇所には下線や記号を付し、具体的な訂正を提示せず、間接的なフィードバックを与えた。生徒は指摘されたエラーについて、「なぜそれがエラーなのか」「どのように直せばよいか」を10分間考察し、口頭または筆記のランゲージングを通じてエラー修正を行った。指導は5ヶ月間で8回実施され、文法テストとテーマ設定をした自由英作文により効果を検証した。

本研究における条件の下では、口頭ランゲージング群は文法テストに効果があった。一方、自由英作文では両群ともに効果は見られなかった。ランゲージングの字数では両群に差はなく、質的分析では、筆記ランゲージング群が気づきのレベルの高いランゲージングを行った。一般的に、ランゲージングの効果検証には、どのような言語技能の学習に導入するか、学習者の英語力、口頭か筆記か、ランゲージングを紙に書くか端末で入力するか、ランゲージングにかかる時間、ランゲージングの頻度と間隔、個別学習か協働学習か、どのような効果検証の方法を採用するか等、様々な要因が関係する。今回は、長期的に口頭と筆記のランゲージングを実践したことで、英語授業にランゲージングを導入するための数々の貴重な示唆を得ることができた点で意義深い。今後は高木氏に、ランゲージングの活用条件を変えて実践を重ね、基礎データを蓄積し、ランゲージングの効果的な活用について研究をさらに深めていただきたい。

B. 実践部門・報告Ⅲ

和泉 伸一

ディスカッション活動において高校生の発話の量と質を高める取り組み

【研究者：久山 慎也】

現代の英語教育における新たな挑戦の一つは、生徒が即興で英語を使ってやり取りをする力をどのように身につけさせるかという課題である。準備が可能な発表活動であれば、考えを練り、原稿を作成し、覚えて練習することで、ある程度見栄えのする発表ができるかもしれない。しかし、即興でのやり取りではそれが難しく、これまでに培ってきた英語力と対話力が試されることになる。その意味でも、生徒の即興でのやり取りの力は、英語教師の日頃の授業力が本当に試される場面と言えるだろう。

久山氏の研究では、このような重要な技能について、高校2年生を対象に2ヶ月間のディスカッション活動指導を行い、その成果を検証している。ディスカッションに有用な定型表現の指導から始まり、具体的な理由を述べる方法、さらに会話を維持・継続するための方略指導を行った。指導前後に行ったディスカッション活動でのやり取りの変化を調査した結果、定型表現の活用によって生徒の英語使用割合が増え、理由づけの量も増加し、限られてはいるが、いくつかのディスカッション方略の出現頻度の増加も見られた。

本研究は実験研究ではないため、統制群のような比較対象は存在しないが、現場で教師が日頃教えている生徒に対して一定期間、系統的な指導を行い、客観的データを収集・分析した、いわば模範的な「アクション・リサーチ」と言えるだろう。この研究から得られる示唆は、実験研究では得難い生態学的妥当性 (ecological validity) を持つと考えられる。今後もタスクや条件を変更し、今回の研究で伸びなかった方略の指導方法を検討するなど、実践に沿った研究の継続を期待している。

B. 実践部門・報告Ⅳ

小泉 利恵

用法基盤モデルに基づいた
スピーキング指導が
中学生の『即興力』育成に及ぼす
効果の検証

【研究者：吉澤 孝幸】

スピーキング力を伸ばすためにどのような指導をすべきか。小中高のそれぞれの文脈で、様々な実践がなされている。教員は授業後に、効果があった、なかったなど感じるだろうが、本研究はその感覚をより客観的な形で検証した。

吉澤氏は「メモ式スピーキング」の効果を調べるために、その指導を受けた実験群と、「リテリング」指導を受けた比較群を設けた。教育的な有効性が先行研究で示されている「リテリング」と、「メモ式スピーキング」を比較することは、二者の違いが見られなかったり、「メモ式スピーキング」が劣ることが示されたりする可能性を考えると、研究デザインに組み込むことは、ある意味勇気が要ったろう。しかし、比較群にも有効な指導を施しながら実験を行い、指導法の優位性を検証することは、教育現場の中で行う研究として望ましいと考える。

本研究の実験群では、スピーキングが直後に向上し、遅延テストでも向上した状態が維持されていた。実験では、事前テスト・直後に行う事後テストの比較がよく用いられるが、長期的な検証のためには、「遅延テスト」も行う方がよい。「事後テスト」と「遅延テスト」の間には検証対象の指導は入れなかったという記載もあり、本研究は丁寧な計画を立てたことがわかる。

また報告書では、実験群・比較群の指導実践の手順が詳細に記載されている。研究では、興味を持った読者が同じように指導できる記述が望ましいが、実際は、読んでも情報不足で再現できない論文も多い。吉澤氏の論文は教員や研究者に非常に参考になる、優れた作品である。

C. 調査部門・報告Ⅰ

竹内 理

日本語母語英語学習者が使用する
定型表現の分析

—「話すこと[やりとり]」と「話すこと[発表]」
の技能育成をめざして—

【研究者：小出 凱渡】

本調査は、英語の話し言葉の2つのモード(モノローグとダイアログ)において、日本人学習者の語連鎖(Lexical Bundles)の使用が、英語母語話者のそれとどう違うのかを、コーパス(ICNALE-Spoken MonologueとICNALE-Spoken Dialogue)を活用して解明しようとしたものである。調査の結果として、日本人英語話者が過剰使用する語連鎖と、過小使用する語連鎖が明らかになり、それぞれの特徴やモード毎の違いなどについても有益な報告がなされている。具体的には、モノローグでは、日本人学習者は特定の句動詞を含む語連鎖や、理由を述べる際に用いる表現を過剰に使用する傾向がみられた。一方で、英語母語話者と比較して、日本人学習者は多様な句動詞表現や助動詞を構成要素とする語連鎖をあまり使用していないことも明らかとなった。また、ダイアログでは、日本人学習者は、同一発話の繰り返しで構成される語連鎖や、いわゆるフィラーを含む表現を多用していたが、多様な時制表現を示す語連鎖やwould like toのような表現は利用できていないことが判明した。

以上のような成果から判断して、本研究はしっかりとした定義に基づき、ユニークな視点から英語学習者の語連鎖使用に切り込んだ力作と言える。今後、得られた結果を実践にどう活かしていくのか、その指導の観点や方法について、中学校での教育経験を積んでいる小出氏が如何に切り込んで行かれるのか、楽しみでしかたがない。

機械翻訳(MT)を学習ツールと するための考察

— 英作文を学ぶ大学生を対象とした量的・質的調査 —

【研究者：湯浅 麻里子】

本研究は、機械翻訳(Machine Translation)を使ってどのようにL2ライティングを行っているかの尺度を図るための学習ツールとするために、実際に英作文を学んでいる大学生を対象とした量的調査と質的調査の両方を実施している。量的調査のリサーチクエスチョンの「MTエンゲージメント尺度はどのような因子で構成されるか」と「MTエンゲージメント尺度の信頼性と妥当性はどの程度か」に対しては、行動・プリエディット認知・ポストエディット認知・情意・社会の5つの因子で構成され、さらに、MTエンゲージメント尺度の信頼性と妥当性は満たされているとの結果が得られている。

質的調査におけるリサーチクエスチョンは、「大学生英語学習者はMTエンゲージメントの5因子の強さによってどのようなグループに分かれるか」「これらのグループのMT自己効力感はどのように異なるか」「各グループの学生はどのようにMTを使っているか」である。低エンゲージメント、中エンゲージメント、高エンゲージメントの3つのグループに分類され、その検証結果が先行研究で言及されていることとほぼ一致していることが認められたという。

しかしながら、MT使用そのものに対する教師の理解度はまだ低いと言わざるを得ない。倫理面への配慮や自己効力感などの心理面も含めて、教育現場とビジネスの現場でのMTの使用実態を比較し、それを教育に活かしていくことが重要である。AIを含めたICTはビジネスにおいて欠くことのできない存在となってきつつあり、近い将来、それが前提となった英語学習に移行していくことは間違いないはずである。

本研究が今後の課題とした限界点は、研究者はもちろん認識する必要があるが、英語教育のみならず社会全体で取り組んでいくべき重要な視点であることも認識しておきたい。

- 第36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	2023年2月1日～4月30日
選 考	2023年5月25日～7月14日
助成金贈呈式	2023年8月14日
研究期間	2023年7月～2024年8月
報告書提出	2024年8月

英語学習者エンゲージメントに関する 文脈的モデルの妥当性検証

研究者: 榎村 祐志 東京都／明治大学大学院 在籍

《研究助言者: 竹内 理》

概要

本研究は、英語学習者におけるエンゲージメントに関する文脈的モデルの妥当性検証を行うことを目的とする実証研究である。関東の中学校、高等学校、大学に通う英語学習者753名を対象に、エンゲージメント、および、その先行要因(心理的欲求、動機づけ)とアウトカム(習熟度)に関する質問紙調査を実施した。予備的分析として因子構造を確認したのち、記述統計にて全体傾向を把握した。次に、多母集団モデルによる共分散構造分析を行った結果、英語学習における心理的欲求・動機づけ・エンゲージメント・習熟度に関する文脈的モデルが妥当であることが示された。最後に、心理ネットワーク分析の結果、変数同士がポジティブに関連しあったシステムを形成しており、中でも自律的動機づけの3変数(内発的、統合的、同一視的)が中心的な役割を果たしていることが明らかになった。

1 はじめに

従来の第二言語習得研究では、様々な個人差要因がある中でも動機づけに注目が集まり、外国語学習の過程において、重要な役割を果たすことが指摘されてきた(Dörnyei & Ushioda, 2021)。これは、自動車の「エンジン」と「ハンドル」に喩えられるように(鹿毛, 2013)、動機づけは長期的なプロセスを要する外国語学習に対して、エネルギーを与え、方向づけるものだからである。しかし、動機づけが高いからといって必ずしも高い学習成果を示すとは限らない。例えば、高く動機づけられた状態で教室に来ていたとしても、英語授業の内容に興味を持てなかったり、高いレベルについていくことができなかったりすることが原因で、積極的に参加できない場面が考えられる。そのため、動機づけだけでなく、実際の学習プロセスを重要視する必要がある。そのような背景から、学習活動における行動に着目した「エンゲージメント(engagement)」に関する研究が近年急増している(Hiniz & Çelik, 2024; Hiver et al., 2024; Namkung & Kim, 2024)。

第二言語習得におけるエンゲージメント研究を包括的に統合したHiver et al. (2024)によれば、(a)エンゲージメントを構成する側面に焦点を当てた研究、(b)エンゲージメントをアウトカムとする研究、(c)エンゲージメントを先行要因とする研究が多く行われている。しかし、これらの要因の関連を包括的に検討した研究は限定的である。そこで本研究は、自己決定理論を援用し、エンゲージメントとその先行要因(心理的欲求・動機づけ)及びアウトカム(習熟度)に関する理論的モデルの妥当性検証を行うことを目的とする。

2 先行研究

2.1 英語学習者エンゲージメント

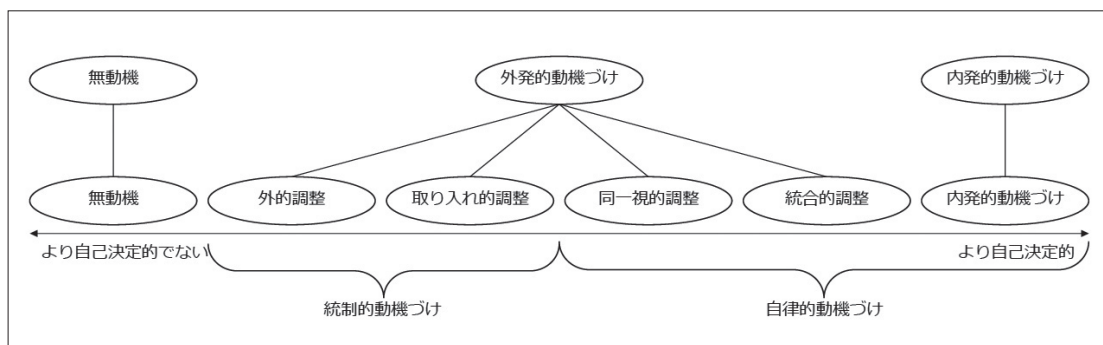
廣森(2024)では、英語学習者エンゲージメントを理解する際、4つの階層と4つの側面から検討する重要性を指摘している。4つの階層は、より高次な階層から、コミュニティ、学校、教室、学習活動で構成されている。学習者は学校内だけでなく学校外での生活もあることから、より広範な意味であるコミュニティという階層がある。次の学校という階層では、日々の学校生活の中で関与する課外活動、学校行事等から検討される。学習者が関与する範囲をより限定した教室という階層では、教師やクラスメイトとの良好な関係性の構築やクラス内のルール遵守といった要因から考えることができる。最も下位の階層だとされる学習活動では、英語授業といった特定の授業科目の中で、積極的に課題に関わっているかどうかという視点からエンゲージメントを捉えている。それぞれの階層は相互作用しながら、エンゲージメントという概念を構成しているとされる。

一方、側面について、言語学習の場面では、行動的・認知的・感情的・社会的の4側面が含まれる。行動的側面は、学習活動における観察可能な行動や積極的参加を意味する。努力、粘り強さ、注意といった指標を用いて測定される。認知的側面は、活動に対する心的努力を意味し、メタ認知、自己調整といった指標から捉えることができる。感情的側面は、学習活動中における情動的な要因(例、興味、楽しさ、意欲)から捉えることが可能である。最後に、ペア・グループ活動が多用される英語授業において重要だと考えられている社会的側面は、学習活動中における他者との関与を意味する。良好な関係性、クラスメイトへの手助けといった指標をもとに考えることができる。これらの側面は個別に成り立っているのではなく、相互作用していると考えられている。そのため、どれか1つだけの側面に注目するのではなく、4側面から包括的に捉えることが重要であるとされる(Hiver et al., 2024)。

以上より、エンゲージメントを捉える際は、どの文脈において、内面的にも外面的にもどの程度積極的に参加しているのかを考慮することが求められる(廣森, 2024)。本研究では、教室レベルのエンゲージメントについて、4側面(行動的・認知的・感情的・社会的側面)から多角的に検討する。

2.2 エンゲージメントの先行要因

エンゲージメントに与える先行要因について、自己決定理論の観点から理解することができる。自己決定理論では、人の行動を自己決定性の程度によって変わるものとして捉えている。自己決定理論は複数のミニ理論を含んでおり、中でも、有機的統合理論(organismic integration theory)は多く援用されているミニ理論の1つである。有機的統合理論は、個人は内在化された動機づけによって行動するだけでなく、他者から影響を受け、様々な価値観や意義を自らに取り入れながら行動する傾向を説明する理論である。この理論の大きな特徴は、動機づけを連続体とみなし、シンプレックス構造(隣接する変数間の関連は、隣接しない変数との関連よりも強くなる性質)が成り立つことである(Howard et al., 2017)。有機的統合理論は、動機づけを大きく3つ(内発的動機づけ、外発的動機づけ、無動機)に分類している(図1参照)。以下では、その3分類について概観していく。



■図1: 有機的統合理論に基づく動機づけ (Howard et al., 2017を参考に著者作成)

最も自己決定性の高い動機づけとして内発的動機づけ (intrinsic motivation) があり、生得的な興味や喜びを得るために行動する状態を指す。次に、外発的動機づけ (extrinsic motivation) がある。この状態では、自身の興味関心からではなく、クラスメイトや教師といった周囲にいる人や、教室の雰囲気をはじめとする学習環境などの外的要因から影響を受けて行動する。外発的動機づけの特徴として、もともと他者や環境からの影響で取り組んでいた学習に対する価値や意義を認識し、徐々に自己決定的に取り組んでいく自律化と呼ばれるプロセスがあるとされている。この自律化の程度によって、外発的動機づけは4つに細分化される。自律化の程度が高いものから順に、統合的調整 (integrated regulation)、同一視的調整 (identified regulation)、取り入れの調整 (introjected regulation)、外的調整 (extrinsic regulation) と位置づけられている。最後に、内発的動機づけ及び外発的動機づけとは別に、無動機 (amotivation) と呼ばれる状態がある。この状態の際、学習者は活動に対する価値を感じず、活動に取り組もうという意図が欠落しているとされる。以上の6分類に加え、自己決定性の程度が高い3種類 (内発的動機づけ、統合的調整、同一視的調整) は、自律的動機づけ (autonomous motivation) と括られる。一方、自己決定性の程度が低い2種類 (取り入れの調整、外的調整) は統制的動機づけ (controlled motivation) と括られている。自己決定性の高い動機づけほど、ポジティブな学習成果をもたらすとされる (Ryan et al., 2022)。

では、どのような要因が動機づけを高めるのか？ この点について、自己決定理論では、基本的心理欲求理論 (basic psychological needs theory) から考えることができる (Ryan et al., 2022)。基本的心理欲求理論は、自律性の欲求 (need for autonomy)、有能性の欲求 (need for competence)、関係性の欲求 (need for relatedness) という3つの心理的欲求から構成されている。自律性の欲求とは、自ら行動を選択しようとする傾向性を指す。有能性の欲求とは、課題を上手くこなすことができると思う傾向性を指す。関係性の欲求とは、他者と上手く関わろうとする傾向性を指す。これら3つの心理的欲求が充足されることにより、動機づけの向上に寄与するとされている (Ryan et al., 2022)。

自己決定理論を援用した様々な先行研究がある中で、Noels et al. (2000) は先駆的な研究である。彼女らは、フランス語を学ぶ大学生を対象に、自己決定理論に基づく動機づけ尺度の開発・妥当性検証を実施した。因子分析の結果をもとに、5つの異なる動機づけ (内発的動機づけ、同一視的調整、取り入れの調整、外的調整、無動機) を測る心理尺度を開発した。彼女らの研究の特徴として、内発的動機づけを、知識 (intrinsic motivation to know)、刺激 (intrinsic motivation to experience stimulation)、達成 (intrinsic motivation to accomplish) の3つの下位要素からなるものと捉えている。しかし、これらの下位要素間の相関係数は非常に高く、下位要素を想定しないことを推奨する研究もある ($r = .86-.96$; Howard et al., 2017)。

自己決定理論は日本の英語学習を対象とする研究にも広く応用されている。廣森 (2003) は日本の高校生275名を対象に、心理的欲求と動機づけ (内発的動機づけ、同一視的調整、外的調整、無動機) の因果関係について検討している。共分散構造分析の結果、3つの心理的欲求の中でも、特に有能性が4つの動機づけ全てを予測した。また、関係性は内発的動機づけと同一視的調整を有意に予測した一方、自律性はどの動機づけにも有意に予測しなかった。

Agawa and Takeuchi (2017) では、大学生を対象に、心理的欲求と動機づけ(内発的動機づけ、同一視的調整、外的調整、無動機)の因果関係を検討している。共分散構造分析の結果、有能性が4つの動機づけを有意に予測した。自律性も内発的動機づけ、同一視的調整、無動機を有意に予測しており、関係性については内発的動機づけのみ予測することが明らかになった。

ここまで言及してきた心理的欲求と動機づけについて、特に有能性や自律性は、学習者の自律的な動機づけを向上させる上で重要な要因であることが窺える。

2.3 エンゲージメントのアウトカム

ここまで、どのような要因がエンゲージメントに影響を与えるのかについて述べた。では、エンゲージメントが高まったことにより、どのようなアウトカムに影響があると考えられるのか。この点について、教育心理学の分野では、Wong et al. (2024) が最も包括的にレビューしている。含まれているアウトカムの指標には、テストスコア・成績・教師による評価といった学習成果に加えて、人生満足度・ポジティブ感情・幸福度といった主観的ウェルビーイングがある。適格性基準を満たした137本の一次研究を対象にメタ分析を実施した結果、行動的、認知的、感情的側面の順に学習成果と正の関連が見られた(順に $r = .39, .31, .26$)。つまり、学習成果を高めるためには、直接観察可能な行動的側面を高めることが重要そうである。興味深いことに、主観的ウェルビーイングとの関連は、学習成果との関連とは逆の順序になっていた。最も関連が強かったのは感情的側面($r = .40$)であり、次いで認知的($r = .35$)、行動的($r = .31$)側面であった。これらの結果から、エンゲージメントのアウトカムは、研究目的によって適切なものを設定することが重要であると示唆される。

2.4 エンゲージメント、先行要因、アウトカムの関連

以上で言及したエンゲージメントとその先行要因・アウトカムの関係について検討した研究は限定的ではあるが、教育心理学の分野ではLam et al. (2012) が大規模な調査を通して、文脈的モデルを提示している。中国の中学生822名を対象に、エンゲージメント(行動的、認知的、感情的)とその先行要因(指導的文脈、社会的関係性、動機づけ要因)およびアウトカム(成績、情緒面、振る舞い)の関連を検討した。相関分析の結果、先行要因の中でも動機づけ要因がエンゲージメントと強い関連が見られた。アウトカムの観点では、ポジティブ感情や成績と関連が見られた。

言語学習の分野では、Dincer et al. (2019) がエンゲージメント(行動的、認知的、感情的、行為主体的)とその先行要因(自律性支援、心理的欲求)およびアウトカム(学習成果、出欠席)の関連を検討している。トルコの英語学習者412名を対象に質問紙調査を行ったのち、共分散構造分析を実施している。分析の結果、心理的欲求は、感情的側面($\beta = .89$)や行為主体的側面($\beta = .83$)を強く予測し、さらに、その2側面は学習成果を正に予測していた(それぞれ、 $\beta = .15/.16$)。

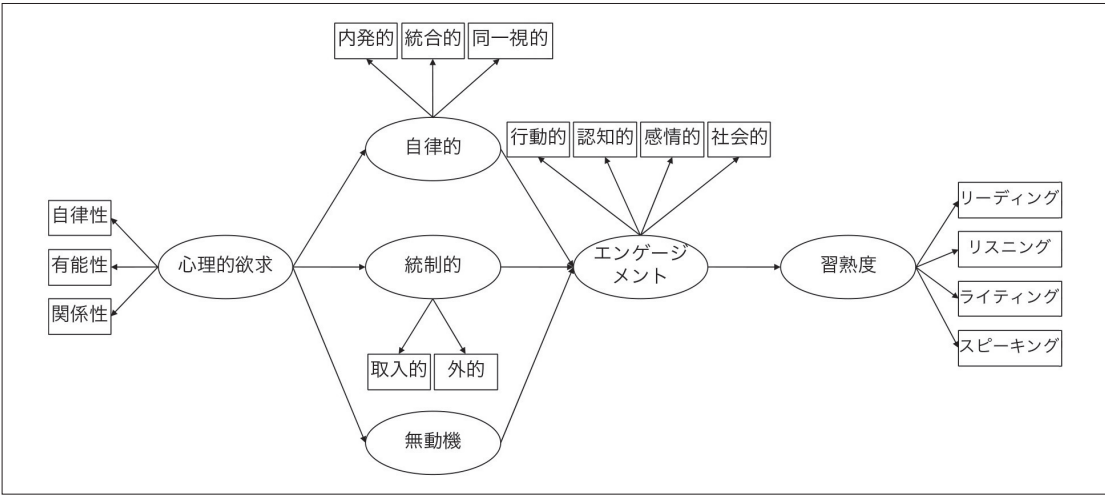
日本の文脈では、Someya and Obermeier (2023) がエンゲージメントとその先行要因(自律性支援、心理的欲求)とアウトカム(テストスコア)の関係について、共分散構造分析を行っている。日本で英語を学ぶ大学生316名を対象に行った調査の結果、心理的欲求がエンゲージメントを強く予測していた($\beta = .86$)。

以上の結果より、自己決定理論を援用した知見が蓄積されつつあり、中でも、心理的欲求がエンゲージメントに正の影響を与えることが示唆されている。

2.5 本研究

以上まで言及したように、エンゲージメントと先行要因、エンゲージメントとアウトカムの関係性に関する個別の知見が蓄積されている。しかし、以下の3点において、限界点があると考えられる。1点目に、文脈的モデルを提唱したLam et al. (2012) では、相関分析を用いているに留まっており、エンゲージメントが

先行要因から受ける影響、エンゲージメントがアウトカムに与える影響について包括的に検討する余地が残されている。本研究では、共分散構造分析を用いてこの課題点を克服する。2点目に、言語学習における文脈において、エンゲージメント・先行要因(心理的欲求、動機づけ)・アウトカムの関係性を包括的に検討した研究は未だ少ない。特に、動機づけとエンゲージメントを同時に検討した研究は限定的である(Namkung & Kim, 2024)。さらに、櫻村(2024)では、学校種による関連性の違いに関する調査の必要性が言及されている。本研究では、自己決定理論を援用して、中高生および大学生の英語学習者のエンゲージメント・先行要因・アウトカムの関連性を、多母集団モデルを用いて検証する。3点目に、従来の先行研究では、共分散構造分析を用いたモデル検証がされてきた。本手法を用いることで変数間の相関・因果関係が明らかにできる一方で、変数同士が複雑に関連し合う中で、介入の影響力を最大限発揮するにはどの要因を優先的に高める必要があるのかを特定するには、方法論的限界点があると考えられる。そこで本研究では、心理ネットワーク分析を用いて含まれる要因間の複雑な関連性を明らかにし、実践者に対して優先的に介入すべき要因に関する教育的示唆を提示することを目的とする。以上の目的を検討するため、次の2つの研究課題を設定した。



■図2: 仮説モデル

- 研究課題 1

英語学習において、エンゲージメント、先行要因(心理的欲求、動機づけ)、アウトカム(習熟度)の関連について、中高生と大学生ではどのような違いが見られるか？
- 研究課題 2

エンゲージメント、先行要因、アウトカムが複雑に関連する中で、中高生と大学生では、それぞれどの要因が中心的な役割を果たしているか？

3 方法

3.1 調査協力者

本研究では、関東の中学校・高校・大学で英語を学ぶ計753名より調査協力を得た。性別は男性190名、女性563名である。校種は中学生32名、高校生371名、大学生350名である。調査協力者には、調査の目的、匿名性の厳守、調査への参加は自由であり、調査途中でも参加をやめることができること、所要時間(15分程度)などを説明した上で、本調査への参加の同意を得た。

3.2 測定具

3.2.1 基本的心理欲求尺度

Leeming and Harris (2022) の基本的心理欲求尺度 (Basic Psychological Needs Satisfaction and Frustration Scale; BPNSFS) を用いて、心理的欲求を測定した。この尺度は、心理的欲求の3要素である自律性(例, 英語を学ぶ中で、私は学ぼうという気持ちが強い)、有能性(例, 英語を学ぶ中で、私は英語に対する自信を持っている)、関係性(例, 英語を学ぶ中で、私はクラスメイトと一緒に頑張っている)を測定する計9項目から構成されている。Leeming and Harris (2022) と同様に、1点(全く当てはまらない)から6点(とても当てはまる)までの6件法で、調査協力者に回答を求めた。

3.2.2 動機づけ尺度

Leeming and Harris (2024) の動機づけ尺度 (Language Learning Orientation Scale; LLOS) を用いて、英語学習に対する動機づけを測定した。この尺度は、内発的動機づけ(例, 私が英語を学ぶ理由は、英語に接すること自体が好きだから)、統合的調整(例, 私が英語を学ぶ理由は、英語を学ぶことが私の一部になっているから)、同一視的調整(例, 私が英語を学ぶ理由は、自分にとって重要な目標を達成することに役立つと思うから)、取り入れ的調整(例, 私が英語を学ぶ理由は、自分には英語を学ぶ義務があると思うから)、外的調整(例, 私が英語を学ぶ理由は、将来より良い仕事に就くため)、無動機(例, 英語を学ぶことについての、私の考えは、英語の学習は時間の無駄であるという感覚がある)を測定する計18項目から構成されている。Leeming and Harris (2022) と同様に、1点(全く当てはまらない)から6点(とても当てはまる)までの6件法で、調査協力者に回答を求めた。

3.2.3 エンゲージメント尺度

行動的・認知的・感情的側面を測定するために Hiver et al. (2020)、社会的側面を測定するために Khajavy (2021) を用いた。行動的(例, 英語の授業では、理解しにくいときでもがんばり続けた)、認知的(例, 英語の授業では、自分が正しくできているか、ていねいに確認した)、感情的(例, 英語の授業では、新しいことを学ぶのが楽しかった)、社会的(例, 英語の授業では、クラスメイトの意見や考えをわかろうとした)を測定する計12項目から構成されている。1点(全く当てはまらない)から6点(とても当てはまる)までの6件法で、調査協力者に回答を求めた。

3.2.4 習熟度

Sudina and Plonsky (2021a, 2021b) を参考に、4スキルに関する自己評価尺度を用いた。本尺度は、リーディング(他の人と比べて、あなたのリーディング力を教えてください)、リスニング(他の人と比べて、あなたのリスニング力を教えてください)、ライティング(他の人と比べて、あなたのライティング力を教えてください)、スピーキング(他の人と比べて、あなたのスピーキング力を教えてください)を測定する計4項目から構成されている。1点(全くできない)から5点(とてもできる)までの5件法で、調査協力者に回答を求めた。

3.3 データ分析

2つの研究課題に関する分析を行う前に、以下の2つの予備的分析を実施した。第1に、欠損値がないことを確認した上で、使用した尺度に関する因子構造を、検証的因子分析(最尤法)を用いて行った。本研究でモデルの評価に用いた指標は、(a) 因子負荷量、(b) χ^2 値、(c) comparative fit index (CFI)、(d) standardized root mean square residual (SRMR)、(e) root-mean-square error of approximation (RMSEA) とその信頼区間(90%)である(In'nami & Koizumi, 2011)。それぞれの指標について、因子負荷量 $\geq .40$ 、CFI $\geq .90$ 、SRMR $\leq .08$ 、RMSEA $\leq .06$ を満たす場合、モデルが適していると判断した(Hu

& Bentler, 1999; Peterson, 2000)。

第2に、以上のモデルの評価に関する結果をもとに、記述統計(平均値、標準偏差、歪度、尖度)及び信頼性係数を算出した。Roever and Phakiti (2018)の基準をもとに、歪度と尖度が ± 2 の範囲内にある場合、データが正規分布していると判断した。

以上の予備的分析をもとに、研究課題1を検討するため、多母集団モデルによる共分散構造分析(最尤法)を実施した。モデルは、中高生と大学生に分けて比較した。モデルの適合度について、検証的因子分析で用いた指標と同じものを採用した。

次に、研究課題2を検討するため、心理ネットワーク分析を行った。これは、収集したデータをもとに変数間の複雑な相互作用を図示する探索的な統計手法である(Freeborn et al., 2023; Isvoranu et al., 2022)。なお、本研究では連続変数で集められた横断データをもとに分析するため、ガウシアン・グラフィカルモデル(gaussian graphical model; GGM)による心理ネットワークを検討する。分析手順は国里他(2023)を参考に、(a)ノード間のエッジの推定とエッジの重み付けの正確度の検討、(b)心理ネットワークの図示、(c)中心性指標の算出及び中心性指標の安定性の検討の3つの手順を踏んだ。はじめに、心理ネットワークのノード間のエッジの推定を行った。用いた推定法は、graphical least absolute shrinkage and selection operator (GLASSO) と Extended Bayesian Information Criterion (EBIC) を合わせた EBICglasso である。この推定法を用いることで、ノード間の関連が弱いエッジを特定し、そのエッジの値をゼロにして心理ネットワーク上から消すことが可能となる。関連が強いエッジのみが図示される結果、より疎(sparse)な心理ネットワークになり、結果の解釈が容易になる。以上の推定法を用いて、心理的欲求と動機づけを含む心理ネットワークを図示した。また、心理ネットワークはサンプルサイズなどをもとに統計的に推定されるものであるため、図示した心理ネットワークのエッジがどの程度正確に重み付けされているかを検討した。Epskamp et al. (2018)の方法に従い、ノンパラメトリック・ブートストラップ法($n = 2,500$)による95%信頼区間を算出した。ブートストラップ法の95%信頼区間が狭いほど、より正確にエッジが推定されていると判断される(Epskamp et al., 2018)。

また、心理ネットワーク内の中心的要因を特定するため、以下の3つの中心性指標(強度、近接性、媒介性)を算出した。強度(strength)は、あるノードが他のノードとどの程度強く関連しあっているかを表す指標である。強度の値が大きいほど、他のノードと強く関連しあっており、他のノードを変化させる影響力が大きいことを意味する。近接性(closeness)は、あるノードが他のノードとどの程度近い場所に位置づけられているかを表す指標である。近接性の値が大きいほど、類似した変数のノードと密集した場所に位置づけられていることを意味する。媒介性(betweenness)は、あるノードがどの程度、他2つのノードの最短経路上にあるのかを表す指標である。媒介性の値が大きいほど、他2つのノードによって媒介される回数が多く、結果として心理ネットワークを変化しづらくさせていることを意味する。これら3つの中心性指標を用いて、心理ネットワークに含まれるノードのうち、どのノードが中核的な役割を果たしているかを特定することができる。中心性指標も、エッジの重み付けと同様に、サンプルサイズなどをもとに統計的に推定されるものであるため、算出された中心性指標はどの程度安定しているかを検討した。Epskamp et al. (2018)に基づき、サンプル数を2,500とするケースドロップ・ブートストラップ法を用いた。この手法は、中心性指標の安定性を検討する目的でEpskamp et al. (2018)により開発された。具体的には、データセットの調査協力者の人数を減らしていった場合(=ケースドロップ)、中心性指標の順序が変わらず、どの程度安定しているかを検討する。各中心性指標の安定性は、correlation stability coefficient (CS係数)によって数値化することができる。中心性指標が安定していると判断するためには、CS係数が.25を超えている必要があり、.50を超えることが望ましいとされる(Epskamp et al., 2018)。

以上の分析は、統計ソフトR(ver. 4.3.0; R Core Team, 2023)を用いて行った。使用したパッケージは、lavaan (ver. 0.6.17; Rosseel, 2012)、psych (ver. 2.3.12; Revelle, 2023)、qgraph (ver. 1.9.8; Epskamp et al., 2012)、bootnet (ver. 1.5.6; Epskamp et al., 2018)、NetworkComparisonTest (ver. 2.2.2; van Borkulo et al., 2023)である。

4 結果

4.1 予備的分析の結果

使用した尺度の因子構造を確認するため、検証的因子分析を行った(表1参照)。基本的心理欲求尺度(BPNSFS)について、RMSEAの値が若干高い値を示していたものの、他の適合度指標は良好であった。動機づけ尺度(LLOS)とエンゲージメント尺度(Engagement)については、全適合度指標において、良好であった。以上の結果より、本研究で使用した尺度は適切な因子構造であると判断し、その後の分析に使用した。

■表1: 検証的因子分析

各尺度	χ^2	CFI	SRMR	RMSEA [90%CI]
BPNSFS	150.64*** (24)	.96	.05	.09 [.07, .10]
LLOS	396.05*** (120)	.96	.04	.06 [.05, .06]
Engagement	172.50*** (48)	.97	.03	.06 [.05, .07]

*** $p < .001$.

次に、使用した下位尺度の記述統計(平均値、標準偏差、歪度、尖度)および信頼性を算出した(表2参照)。平均値の観点から、本研究の調査協力者はおおよそ外発的に動機づけられた学習者であり、行動的・認知的に英語学習に取り組んでいることが窺える。また、全ての下位尺度の歪度と尖度が±2の範囲内に収まっていたことから、データは正規分布していると判断した。最後に、下位尺度の信頼性係数について、取り入れ的调整が.67と比較的低い値を示していたものの、その他の尺度は $\alpha = .77-.88$ であり、内的一貫性が保たれていると判断した。

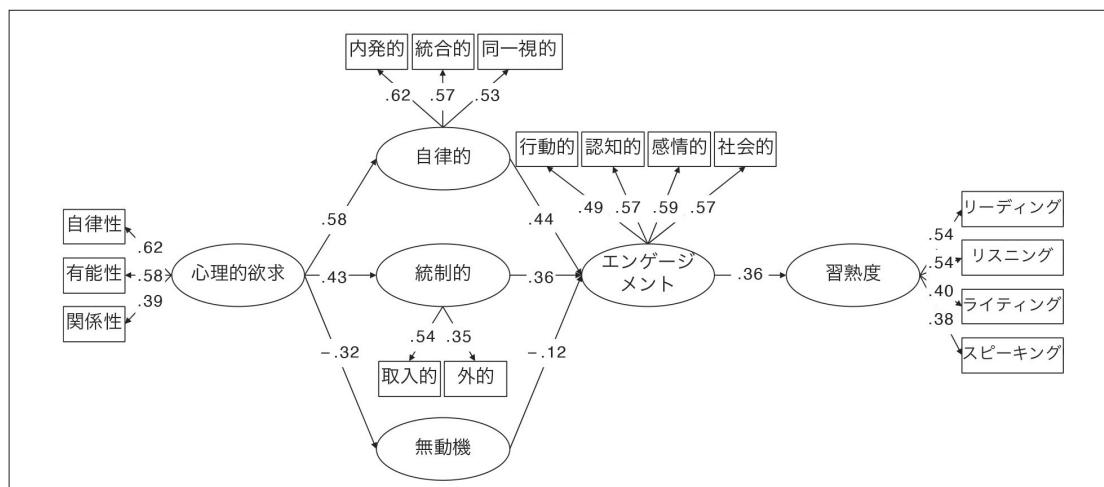
■表2: 記述統計(中高生)

変数	M	SD	歪度	尖度	α	95% CI
自律性	3.64	1.13	0.11	-0.36	.85	[.83, .87]
有能性	3.18	1.17	0.18	-0.36	.84	[.81, .86]
関係性	3.65	1.15	-0.18	-0.33	.84	[.82, .86]
内発的	3.46	1.29	0.00	-0.44	.88	[.87, .90]
統合的	3.10	1.24	0.16	-0.38	.87	[.85, .88]
同一視的	3.75	1.22	-0.20	-0.44	.84	[.82, .86]
取り入れ的	3.62	1.05	0.08	-0.03	.67	[.62, .71]
外的	4.06	1.10	-0.19	-0.18	.78	[.75, .81]
無動機	2.83	1.28	0.35	-0.57	.87	[.86, .89]
行動的	3.81	1.09	-0.13	-0.08	.84	[.82, .86]
認知的	3.80	1.05	0.06	-0.28	.81	[.79, .83]
感情的	3.50	1.24	-0.06	-0.41	.88	[.86, .90]
社会的	3.72	1.12	-0.21	-0.09	.77	[.74, .80]
習熟度	3.03	0.93	-0.26	-0.09	.84	[.82, .86]

注. CI = 信頼区間。

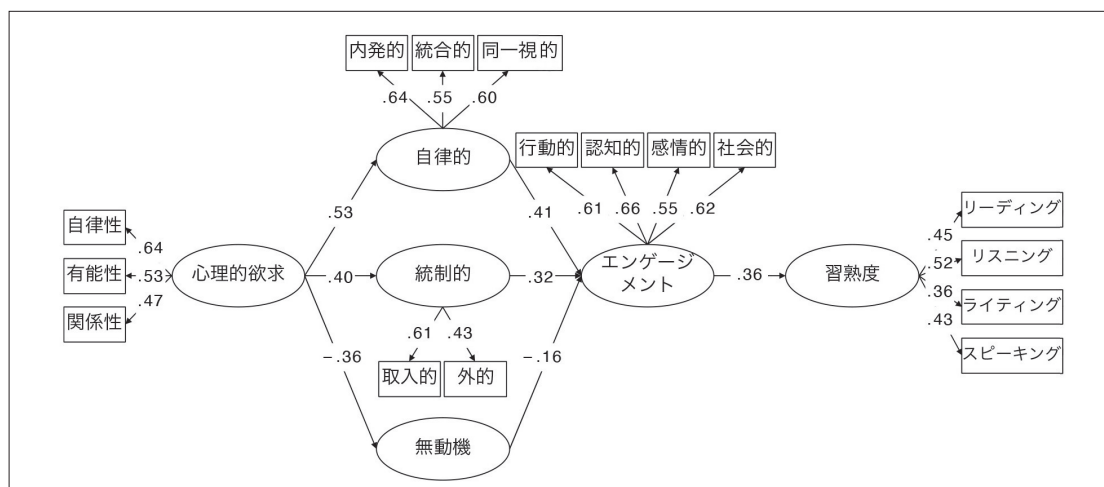
4.2 研究課題 1

エンゲージメントと先行要因(心理的欲求、動機づけ)およびアウトカム(習熟度)の関連において、中高生と大学生で異なるのかを明らかにするため、多母集団モデルによる共分散構造分析を行った。はじめに、中高生を対象にしたモデルの結果を図3に示す。モデルの適合度指標について、 $\chi^2(841) = 2348.58 (p < .001)$, CFI = .89, SRMR = .06, RMSEA = .07 [.06, .07]であり、良好な値を示していた。各変数間の関連について、心理的欲求は自律的・統制的動機づけを正に予測しており($\beta = .58/.43$)、自律的・統制的動機づけもエンゲージメントを正に予測していた($\beta = .44/.36$)。最後に、エンゲージメントは習熟度に対して正の予測を示していた($\beta = .36$)。



■図3: 共分散構造分析の結果(中学生・高校生) 注. 値は標準化係数(全て $p < .05$)。

次に、大学生を対象にしたモデルの結果を図4に示す。適合度指標について、 $\chi^2(841) = 2032.75 (p < .001)$, CFI = .87, SRMR = .08, RMSEA = .07 [.06, .07]であり、良好な値を示していた。各変数間の関連は中高生のモデルと類似した結果が得られた。心理的欲求は自律的・統制的動機づけを、自律的・統制的動機づけはエンゲージメントを、そしてエンゲージメントは習熟度を正に予測していた。

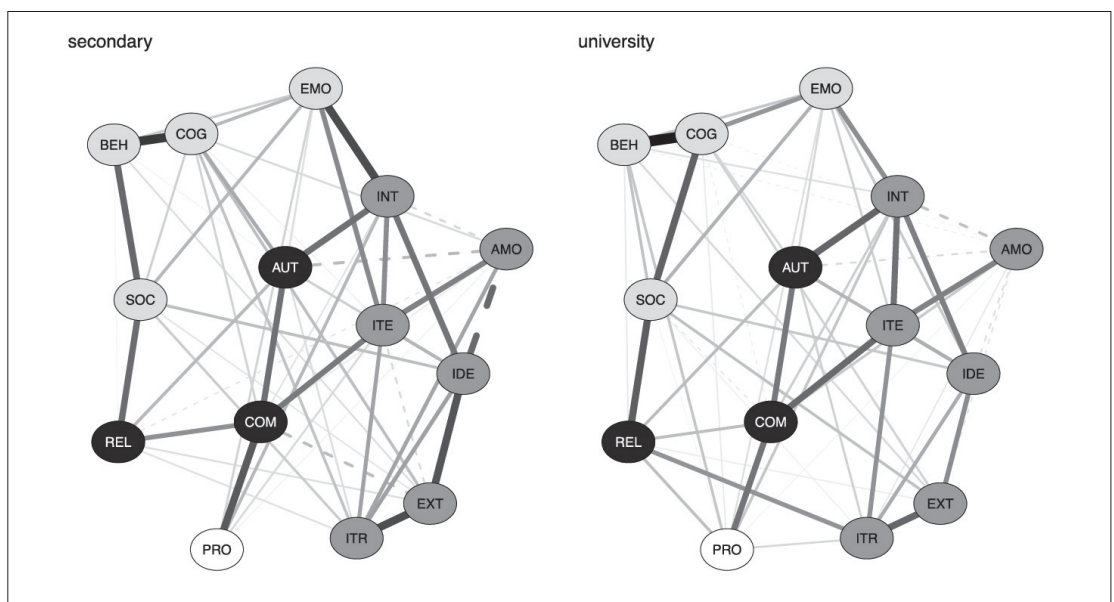


■図4: 共分散構造分析の結果(大学生) 注. 値は標準化係数(全て $p < .05$)。

以上の結果をまとめると、中高生・大学生問わず、心理的欲求は英語学習に対する動機づけを高める要因であると言える。さらに、動機づけを高めることでエンゲージメント促進に繋げることができると考えられる。最終的には、エンゲージメントが高まることで、習熟度の向上に寄与していくことが示唆された。

4.3 研究課題 2

次に、心理的欲求、動機づけ、エンゲージメント、アウトカムが複雑に関連しあう中で、中心的な役割を果たす要因を特定するため、心理ネットワーク分析を実施した。また、ネットワーク比較検定を通して、ネットワーク構造および中心性指標の相違点・類似点を検討した。それぞれのネットワーク構造を図5に示す。なお、エッジの重み付けの正確性を検討したところ(資料1および2参照)、本研究で得られたエッジの値とブートストラップ法による値が重複しており、信頼区間も比較的狭かったことから、比較的正確に重み付けされていると判断した。



■図5: 心理ネットワーク分析の結果

注. 実線のエッジは正の偏相関関係, 点線のエッジは負の偏相関関係を表す。左は中高生, 右は大学生の心理ネットワーク図である。

AUT = 自律性, COM = 有能性, REL = 関係性, INT = 内発的動機づけ, ITE = 統合的調整, IDE = 同一視的調整, ITR = 取り入れの調整, EXT = 外的調整, AMO = 無動機, BEH = 行動的, COG = 認知的, EMO = 感情的, SOC = 社会的, PRO = 習熟度。

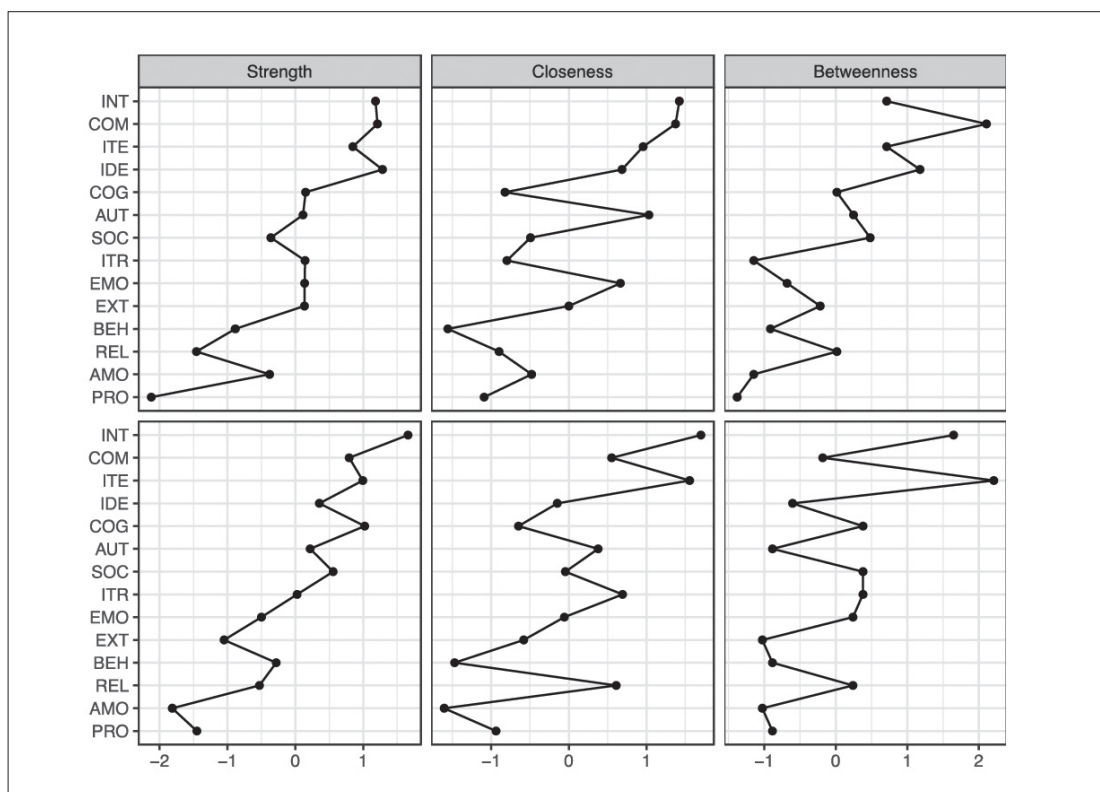
中高生・大学生に共通して、動機づけ間、エンゲージメント間で強い偏相関関係があることが明らかになった。特に、どちらのネットワークにおいても、行動的側面と認知的側面が強く関連しあっていることがわかる。また、基本的には正の偏相関関係が見られる一方、無動機が他変数と負の関連があることも共通してみられる。異なる点として、エンゲージメントと習熟度の関連が挙げられる。中高生の場合、習熟度と関連が見られるのは感情的側面のみである一方、大学生の場合、4側面(行動的, 認知的, 感情的, 社会的)全てで関連が見られた。

次に、中心的要因を特定するため、3つの中心性指標を算出した。また、中心性の安定性を検討するため、CS係数を算出した。中高生の心理ネットワークでは、強度(Strength)/近接性(Closeness)/媒介性(Betweenness)それぞれ、.59/.44/.28であった。同様に、大学生の心理ネットワークでは、.67/.36/.20であっ

た。以上より、どちらの中心性指標でも、強度と近接性の結果は安定していると判断できる一方、大学生の心理ネットワークにおける媒介性は基準値である.25を下回っていたため、判断には注意が必要であると考えられる。

以上の結果をもとに、図6に示す中心性指標を見てみると、自律的動機づけとされる内発的・統一的・同一視的はどちらのネットワークでも比較的中心的な役割を果たしていることがわかる。エンゲージメントの中では、認知的側面が中心的な役割を果たしていると考えられる一方、行動的側面はあまり中心的な役割を果たす要因ではなさそうである。また、図5で他変数と負の関連を示していた外的調整や無動機はあまり中心的な役割を果たしていないことが窺える。

以上の心理ネットワーク分析に関する一連の結果より、基本的に、心理的欲求、動機づけ、エンゲージメント、習熟度は正に相互作用しあっているシステムとして構成されており、ある変数がポジティブに変わると、その影響が伝染して他変数もポジティブに変化する好循環を生み出している特徴を持っていると考えられる。



■図6: 中心性指標

注. 上段は中学生・高校生, 下段は大学生における心理ネットワークの中心性指標。縦軸は各ノードを表し、横軸はz値を表す。
AUT = 自律性, COM = 有能性, REL = 関係性, INT = 内発的動機づけ, ITE = 統一的調整, IDE = 同一視的調整, ITR = 取り
入りの調整, EXT = 外的調整, AMO = 無動機, BEH = 行動的, COG = 認知的, EMO = 感情的, SOC = 社会的, PRO = 習熟度。

5 考察

本研究では、これまで第二言語習得研究の分野であまり検討されてこなかったエンゲージメントとその先行要因およびアウトカムの関連について、自己決定理論を援用して包括的に調査した。共分散

構造分析と心理ネットワーク分析を併用して分析した結果、心理的欲求→動機づけ→エンゲージメント→習熟度の順に正に予測していた。また、それぞれポジティブに関連しあったシステムとしての特徴を持つことが明らかになった。

研究課題1に関する心理的欲求と動機づけの関連について、日本の高校生を対象に調査した廣森(2003)、及び日本の大学生を対象にしたAgawa and Takeuchi(2017)を支持する結果であった。つまり、3つの心理的欲求が充足されることで、動機づけにポジティブな影響を与えるという自己決定理論の理論的前提をより強固なものにしている。さらに、本研究ではエンゲージメントを含め、心理的欲求・動機づけ・エンゲージメントの関連を検討した。この3者を同時に検討した先行研究は少ないが、上述した廣森(2003)とAgawa and Takeuchi(2017)に加えて、心理的欲求とエンゲージメントの関連を扱っているDincer et al.(2019)、Someya and Obermeier(2023)を補完する結果であった。これまでの先行研究では、心理的欲求が動機づけ、またはエンゲージメントを正に予測することが指摘されてきた。以上の知見に加え、本研究の結果は、動機づけが心理的欲求とエンゲージメントを媒介する要因であることを実証的に示している。言い換えれば、英語学習に対する心理的欲求が充足されると、より自律的な動機づけを有し、学習活動に積極的に関与するようになることを示唆している。

エンゲージメントとアウトカムの関連については、Dincer et al.(2019)、Someya and Obermeier(2023)と同様に、正の影響が見られた。ただし、本研究はより強くアウトカムを予測している結果であった。これは、扱っている指標に起因するものだと考えられる。Dincer et al.(2019)、Someya and Obermeier(2023)では、テストスコアをエンゲージメントのアウトカムとして設定している一方で、本研究では、自己評価による習熟度を採用した。そのため、今後、客観的指標を含む多様なアウトカムを用いて、エンゲージメント向上によってもたらされる影響を検討することが必要だと考える。

研究課題2について、心理ネットワーク分析の結果(図5)と中心性指標(図6)の結果より、自律的動機づけと括られる内発的動機づけ、統合的調整、同一視的調整が中核的な役割を果たしていることが示唆された。特に、これら3変数は、統制的動機づけ間より強い相関関係にあるというメタ分析(Howard et al., 2017, 2021)の結果を支持するものである。つまり、自律的動機づけが高まることにより、効率的に他の周辺変数に影響を与え、ポジティブな変化を生み出すことができると考えられる。

一方、ネットワーク図では、外的調整と無動機は他変数と負の関連があることが見られた。この点は、教育心理学や第二言語習得研究の先行研究の知見を支持するものである(例、Howard et al., 2021; Leeming & Harris, 2024)。つまり、日本で英語を学ぶ学習者にとって、英語を学ぶ目的や意思が欠けてしまうと他の調整段階である動機づけを阻害してしまう可能性を示唆している。本研究では、これらの先行研究の知見に対して、ネットワーク図と中心性指標より新しい知見を提供している。それは、他変数と負の関連があるが、中心的な役割ではなく、負の影響力が弱いという点である。この点について、本研究で検討した心理ネットワークでは変数のほとんどが正の関連を示しており、より正の影響力が強いことが理由だと考えられる。つまり、たとえ英語学習に対する目的を見失ったり、他者からの影響で勉強させられていると感じていたりしても、そういった負の影響力は弱い。むしろ、心理的欲求を満たし自己決定性の高い動機づけを有することで、エンゲージメントを促進することができることを示唆している。

6 結論

本研究では、以下の3点で限界点があると言える。はじめに、本研究では、教室レベルのエンゲージメントを検討しているが、先行研究で言及したようにエンゲージメントは複数の階層から捉えることができる。例えば、教室レベルのエンゲージメントが高いとしても、タスクレベルでは高くない可能性もある。そのため、本研究で検証したモデルが他のレベルでも妥当なモデルかどうかの検証が求められる。2点目

に、本研究では、エンゲージメントのアウトカムの指標として、自己評価による習熟度を用いた。学習者のエンゲージメントとアウトカムの関連について、先行研究を包括的に統合した Wong et al. (2024) では、テストスコア等の客観的指標や主観的幸福感といった指標も含めて包括的に検討している。今後、様々な指標を用いて検討することで、エンゲージメントの影響を多角的に理解することが求められる。3点目に、本研究で得られた心理ネットワーク図(図5)は探索的手法によって得られたものであるため、検証的手法を用いて心理ネットワークの構造を検討することが必要になる。

以上の限界点はあるものの、英語学習者のエンゲージメントとその先行要因・アウトカムを含めて包括的な関係性を検討した。今後、文脈的モデルや心理ネットワーク構造を繰り返し検証していくことで、優先して介入すべき要因に関するより頑強な知見を得ることができると考える。そのような研究が蓄積されることで、ネットワーク全体を効率的にポジティブに変化させる具体的な示唆を提示することが期待される。

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、および選考委員の皆様にご心より感謝申し上げます。特に、研究助言者である関西大学の竹内理先生には研究実施に際して貴重なアドバイスを頂戴いたしました。深く感謝申し上げます。また、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明治大学の廣森友人先生、および同大学院ゼミの皆様にご深く感謝申し上げます。最後に、本研究にご協力してくださいました中学生・高校生・大学生の皆様、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。

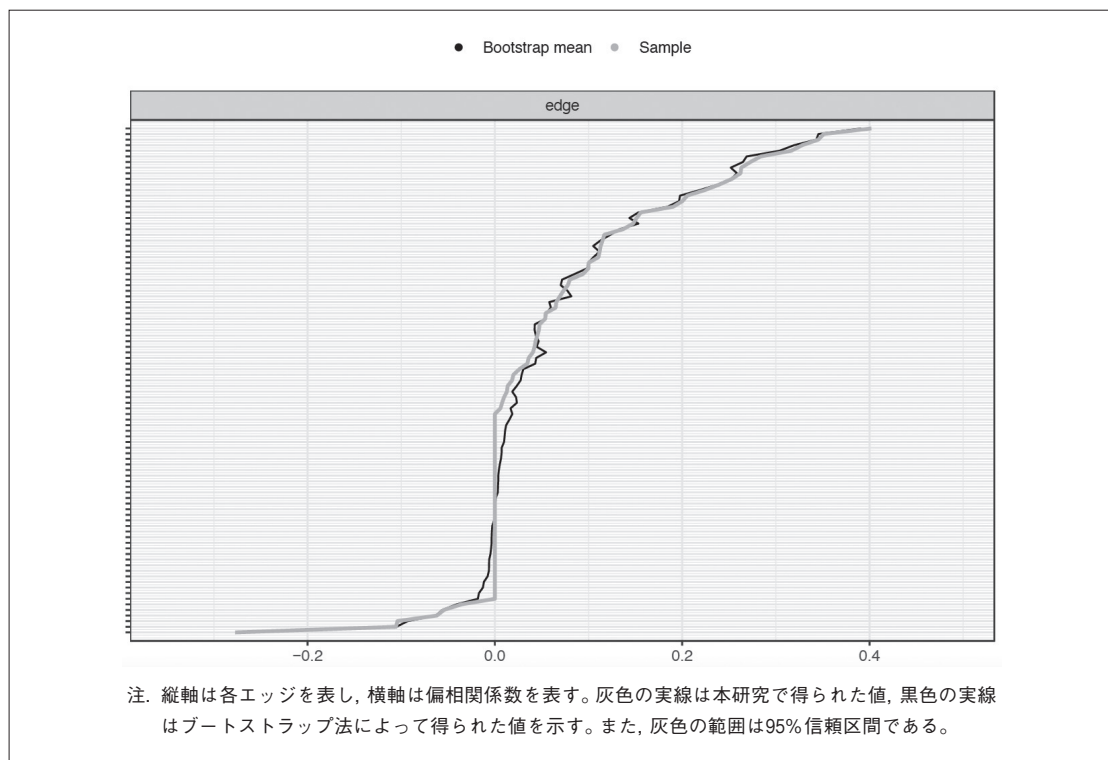
引用文献

- Agawa, T., & Takeuchi, O. (2017). Examining the validation of a newly developed motivation questionnaire: Applying self-determination theory in Japanese university EFL context. *JACET Journal*, 61, 1-21. https://doi.org/10.32234/jacetjournal.61.0_1
- Dincer, A., Yeşilyurt, S., Noels, K. A., & Vargas Lascano, D. I. (2019). Self-determination and classroom engagement of EFL learners: A mixed methods study of the self-system model of motivational development. *SAGE Open*, 9(2), 1-15. <https://doi.org/10.1177/2158244019853913>
- Dörnyei, Z., & Ushioda, E. (2021). *Teaching and researching motivation* (3rd ed.). Routledge.
- Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50(1), 195-212. <https://doi.org/10.3758/s13428-017-0862-1>
- Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(4), 1-18. <https://doi.org/10.18637/jss.v048.i04>
- Freeborn, L., Andringa, S., Lunansky, G., & Rispens, J. (2023). Network analysis for modeling complex systems in SLA research. *Studies in Second Language Acquisition*, 45(2), 526-557. <https://doi.org/10.1017/S0272263122000407>
- Hinüz, G., & Çelik, Ö. (2024). A bibliometric and content analysis of student engagement research in English language teaching. *Language Awareness*. Advance online publication. <https://doi.org/10.1080/09658416.2024.2367970>
- 廣森友人 (2003) 「学習者の動機づけは何によって高まるのか—自己決定理論による高校生英語学習者の動機づけの検討—」 *JALT Journal*, 25(2), 173-186. <https://doi.org/10.37546/JALTJJ25.2-3>
- 廣森友人 (2024) 「エンゲージメントとは何か？」廣森友人・和田玲 (編著) 『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 (pp. 2-13)』大修館書店。
- 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>
- Hiver, P., Zhou, S. A., Tahmouresi, S., Sang, Y., & Papi, M. (2020). Why stories matter: Exploring learner engagement and metacognition through narratives of the L2 learning experience. *System*, 91, 102260. <https://doi.org/10.1016/j.system.2020.102260>
- Howard, J. L., Bureau, J. S., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science*, 16(6), 1300-1323. <https://doi.org/10.1177/1745691620966789>
- Howard, J. L., Gagné, M., & Bureau, J. S. (2017). Testing a continuum structure of self-determined motivation: A meta-

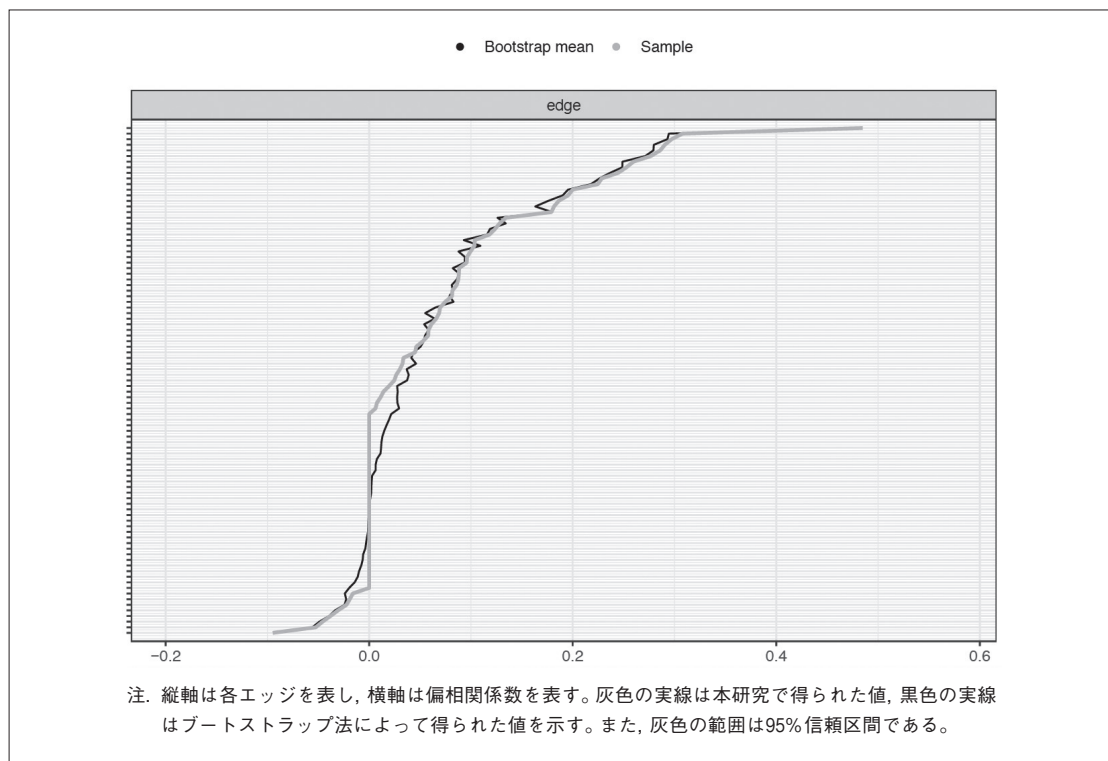
引用文献

- analysis. *Psychological Bulletin*, 143(12), 1346-1377. <https://doi.org/10.1037/bul0000125>
- 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>
- In'nami, Y., & Koizumi, R. (2011). Structural equation modeling in language testing and learning research. *Language Assessment Quarterly*, 8(3), 250-276. <https://doi.org/10.1080/15434303.2011.582203>
- Isvoranu, A.-M., Epskamp, S., Waldorp, L. J., & Borsboom, D. (Eds.). (2022). *Network psychometrics with R: A guide for behavioral and social scientists*. Routledge.
- 鹿毛雅治 (2013) 『学習意欲の理論—動機づけの教育心理学—』金子書房.
- 樫村祐志 (2024) 「何がエンゲージメントを高め、エンゲージメントは何をもたらすのか？」廣森友人・和田玲 (編著) 『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 (pp. 14-24)』大修館書店.
- 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.
- 国里愛彦・藁科佳奈・隅田昌孝・村口ゆう・大水拓海 (2023) 「横断データに対する心理ネットワーク分析のチュートリアル」『専修人間科学論集心理学篇』13(1), 51-67. <https://doi.org/10.34360/00013449>
- Lam, S.-F., Wong, B. P. H., Yang, H., & Liu, Y. (2012). Understanding student engagement with a contextual model. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 403-420). Springer.
- Leeming, P., & Harris, J. (2022). Measuring foreign language students' self-determination: A Rasch validation study. *Language Learning*, 72(3), 646-694. <https://doi.org/10.1111/lang.12496>
- Leeming, P., & Harris, J. (2024). The language learning orientations scale and language learners' motivation in Japan: A partial replication study. *Research Methods in Applied Linguistics*, 3(1), 10096. <https://doi.org/10.1016/j.rmal.2024.100096>
- Namkung, Y., & Kim, Y. (2024). A review of learner motivation and engagement research in task-based language teaching. In S. Li (Ed.), *Individual differences and task-based language teaching* (pp. 198-227). John Benjamins.
- Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, 50(1), 57-85. <https://doi.org/10.1111/0023-8333.00111>
- Peterson, R. A. (2000). A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters*, 11(3), 261-275. <https://doi.org/10.1023/A:1008191211004>
- R Core Team. (2023). *R: A language and environment for statistical computing* (Version 4.3.0) [Computer software]. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University, Evanston, Illinois.
- Roeber, C., & Phakiti, A. (2018). *Quantitative methods for second language research: A problem-solving approach*. Routledge.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Ryan, R. M., Duineveld, J. J., Di Domenico, S. I., Ryan, W. S., Steward, B. A., & Bradshaw, E. L. (2022). We know this much is (meta-analytically) true: A meta-review of meta-analytic findings evaluating self-determination theory. *Psychological Bulletin*, 148(11-12), 813-842. <https://doi.org/10.1037/bul0000385>
- Someya, F., & Obermeier, A. (2023). Autonomy support, psychological needs satisfaction, academic engagement, and achievement in English learning: Generating a mediation model. *LET Kansai Chapter Collected Papers*, 21, 19-38. https://doi.org/10.50924/letkansai.21.0_19
- Sudina, E., & Plonsky, L. (2021a). Academic perseverance in foreign language learning: An investigation of language-specific grit and its conceptual correlates. *The Modern Language Journal*, 105(4), 829-857. <https://doi.org/10.1111/modl.12738>
- Sudina, E., & Plonsky, L. (2021b). Language learning grit, achievement, and anxiety among L2 and L3 learners in Russia. *ITL - International Journal of Applied Linguistics*, 172(2), 161-198. <https://doi.org/10.1075/itl.20001.sud>
- van Borkulo, C. D., van Bork, R., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2023). Comparing network structures on three aspects: A permutation test. *Psychological Methods*, 28(6), 1273-1285. <https://doi.org/10.1037/met0000476>
- Wong, Z. Y., Liem, G. A. D., Chan, M., & Datu, J. A. D. (2024). Student engagement and its association with academic achievement and subjective well-being: A systematic review and meta-analysis. *Journal of Educational Psychology*, 116(1), 48-75. <https://doi.org/10.1037/edu0000833>

資料1: エッジの重み付けの正確度(中学生・高校生)



資料2: エッジの重み付けの正確度(大学生)



第36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

結束性指標に基づく日本人英語学習者の エッセイライティングスコアの予測

研究者:久保 佑輔 茨城県／筑波大学大学院 在籍

《研究助言者:寺内 一》

概要

ライティングタスクでは、語彙や文法の正確性に加えて、テキストの一貫性や論理性の観点からの指導や評価も求められる。このようなテキストのまとまりを測定する指標の1つに結束性が挙げられる。しかし、どのような結束性の特徴がライティングスコアに寄与するのかは十分に明らかになっていない。そこで本研究では、日本人英語学習者の作成したライティングスコアの予測に寄与する結束性指標を調査した。具体的には、テキストの言語的特徴に基づいて結束性指標を算出するTool for the Automatic Analysis of Cohesion(TAACO)を用いて、テキスト一貫性の判断とエッセイの質を予測する変数の特定を行った。その結果、テキスト一貫性とエッセイの質の両方において、内容語の重複がマイナスの影響を与えることが示された。また、3人称代名詞を使用して同じ名詞の過剰使用を避けることで、エッセイの質が高まることも確認された。以上のことから、テキストの一貫性やエッセイの質を高めるためには、隣接文間で同じ内容語の繰り返しの使用を防ぐための指導やフィードバックをする必要性などが示唆された。

1 はじめに

グローバル化が急速に進展している現在では、外国語によるコミュニケーション能力は様々な場面で益々必要となることが想定される。このような背景から、英語による発信能力の強化を目指して、高等学校では新たに「論理・表現」が新設された。この科目では、情報や自分の考えについて論理の構成や展開を工夫しながら話したり書いたりする能力の育成を目標としている(文部科学省, 2018)。つまり、産出した英文の正確性の向上だけでなく、内容を効果的に伝えるためにまとまりのあるプロダクトを産出する能力が求められていると言える。また、英語外部試験(e.g., 英検, IELTS)においても、語彙や文法の正確性に加えて、英文のつながりやまとまりの観点が評価項目に含まれている。以上のことから、プロダクトの論理性や一貫性に焦点を当てた指導や評価も取り入れていく必要がある。

このようなプロダクトの内容や構成のつながりを測定する指標の1つに結束性が挙げられる(Crossley et al., 2016a)。一般的に、質の高いプロダクトにとって結束性は重要な要素とされる(Crossley, 2020)。そのため英語熟達度の高い学習者ほど、分かりやすく伝えるために論理的で一貫したプロダクトの産出が可能になると考えられる。しかし、どのような結束性の特徴が質の高いプロダクトと判断されるのか十分に明らかになっていない(e.g., Crossley et al., 2016b; Tsunemoto & Trofimovich, 2024)。

そこで本研究では、日本人英語学習者の作成したエッセイライティングに焦点を当て、テキスト一貫性やエッセイの質の評価に寄与する結束性指標を明らかにする。結束性指標の算出には、英文自動解析ツ-

ルである Tool for the Automatic Analysis of Cohesion 2.0 (TAACO; Crossley et al., 2019) を使用する。本研究の結果から、ライティング活動における結束性の観点からの指導や評価の方法について示唆を得ることを目指す。

2 先行研究

2.1 結束性

結束性とは、読み手がテキスト内のアイデア同士を結びつけるための明示的な手がかりと定義される (Crossley et al., 2016a)。具体的な明示的な手がかりとして、隣接文間での単語や概念の重複が挙げられる。また、連結詞 (e.g., because, therefore) を使用することで、内容のつながりを読み手に明示的に伝えることもできる (Halliday & Hasan, 1976)。一貫性もテキストのまとまりを示す言語的特性であるが、これは読み手によるテキスト内容の理解度を指す (O’ Reilly & McNamara, 2007)。そのため、背景知識や言語熟達度などの読み手の要因によって一貫性は変化する (McNamara & Kintsch, 1996)。

結束性も一貫性もテキスト内のつながりやまとまりの程度を示す要素である。しかし、一貫性は読み手の主観的な判断を含むため、客観的な指標の算出は困難である。その一方で結束性は、代名詞や連結詞などの表面的な言語的特徴の有無によって変化する (保田, 2024)。したがって、テキスト内のレマの重複の程度や接続語の使用頻度などに基づいて客観的に数値化することができる (Crossley et al., 2016a)。そのため、テキストの結束性を測定する Coh-Metrix (Graesser et al., 2004) や TAACO (Crossley et al., 2016a) などの自動解析ツールがこれまでに開発されている。

結束性の特徴が確認できる英文例を表1にまとめる。(例1) では、前文の college students を代名詞の they で置き換えることによって、隣接文間での結束性を強めている。(例2) では、money という同じ内容を繰り返し使用することによって、隣接文間での結束性を強めている。(例3) では、副詞の lastly を使用することによって、テキスト全体での結束性を強めている。英文自動解析ツールでは、このように結束性指標を数値化している。より詳細な指標の算出方法については Crossley et al. (2019) を参照されたい。

表1: 結束性の特徴が確認できる英文例

結束性の特徴のある英文の抜粋
(例1) And the college students have much time. So if they want
(例2) they can make money by doing part-time job for their own. Planning how to use money is helpful
(例3) Lastly, they have responsibility for their job.

注. ICNALE Edited Essay (Ishikawa, 2018a) の W_JPN_PTJ0_027_A2_0 が作成したエッセイから抽出。結束性の特徴の見られる単語は太字で表示。

2.2 ライティング評価と結束性の関係

評価者によるライティングの質と一貫性の判断との間には関連があることが報告されている (Bae et al., 2016; Crossley & McNamara, 2010; Plakans & Gebriel, 2017)。したがってライティング熟達度の向上には、まとまりやつながりのあるテキストを産出する能力の育成が必要と言える。そのため、ライティング熟達度の予測因子の1つとして、結束性が挙げられている (Crossley, 2020; Kojima & Kaneta, 2022; McNamara et al., 2010)。しかし、どのような結束性指標がライティング評価に影響を及ぼすのかについて、L1とL2の学習者を対象とした両先行研究の間で結果は一致していない (e.g., Crossley et al., 2016b)。

2.2.1 L1ライティングにおける結束性の影響

英語母語話者のライティング熟達度に対する結束性の影響は、学年によって異なる傾向が見られる。初期段階では、隣接文間のつながりを示す局所的結束性がライティングの質に寄与する結果が支持されている (Crossley, 2020)。例えば、小学生が産出する質の高いライティングには、因果関係や対立関係を示す接続語や、代名詞や指示詞が多く含まれることが報告されている (Cameron et al., 1995; Myhill, 2008)。その後、学年が上がるにつれて明示的な結束性マーカーの使用頻度が減っていき、局所的結束性指標はライティングスコアの予測に寄与しないことが確認されている (e.g., McNamara et al., 2010)。その一方で、指示詞の使用がライティングスコアを予測する結果も報告されている (MacArthur et al., 2019)。大局的結束性については、段落間の意味的重複や語彙的重複の指標はライティングスコアの予測に寄与することが確認されている (e.g., McNamara et al., 2013)。

まとめると、ライティングの発達初期段階では、局所的結束性の使用はライティングの質に影響を与える。しかし、ある程度のライティングスキルが身に付くにつれて、局所的結束性の影響はライティングトピックや熟達度によって異なることが推察される。しかし、段落間やテキスト全体での結束性(大局的結束性、テキスト全体の結束性)はライティングスコアと関連する傾向にある。

2.2.2 L2ライティングにおける結束性の影響

L1と同様にL2ライティングにおいても、局所的結束性は一般的にライティングの質にマイナスの影響を与える傾向にある (e.g., Crossley, 2020)。例えば、高校生のL2学習者を対象に検証した Crossley and McNamara (2012) は、内容語の重複が評価にマイナスの影響を与えることを報告している。語彙の重複とライティングの質やテキスト一貫性の判断との間には負の相関があることは、その他の先行研究においても確認されている (e.g., Grant & Ginther, 2000; Guo et al., 2013)。同様に、代名詞の使用もライティング能力との間に負の相関が確認されている (Crossley et al., 2016b)。したがって、ライティング能力の高いL2学習者は、より多様な語彙を用いてエッセイを作成する傾向にあると言える。しかし、多様な語彙の使用は評価結果と負の相関を確認した結果も報告されている (Tywoniw & Crossley, 2019)。

連結詞について、質の高いライティングほど多く使用されることが明らかになっている (e.g., Connor, 1990; Duggleby et al., 2016; Tywoniw & Crossley, 2019)。しかし、等位接続詞の使用は負の予測因子として確認されている (Crossley et al., 2016b)。また、連結詞の使用がスコアの予測に寄与する変数として抽出されていない先行研究もある (e.g., Crossley & McNamara, 2012)。

このように、L2ライティング評価における結束性の影響は十分に明らかになっていない。これは、ライティングにおける言語的特徴は、学習者の要因だけでなく、ライティングタスクによっても異なるためであると推察される (Crossley, 2020; Plakans, 2008; Plakans & Gebril, 2013)。したがって、先行研究から得られた知見を一般化するのは難しく、それぞれの文脈に合わせた検証をすることで、結束性がライティング評価に与える影響を解明していく必要があるだろう。

3 本研究

3.1 本研究の目的

結束性は、ライティングの質と一貫性の判断に影響を及ぼす言語的特徴とされる (Crossley, 2020)。そのため、英文自動解析ツールによって、ライティング評価における結束性の影響について包括的検証がこれまでに実施されている (e.g., Crossley et al., 2016b; Kim, 2022; Tywoniw & Crossley, 2019)。しかし、学習者やタスクの要因が影響することから、先行研究間で報告されている結果は一致していないため、特定の文脈ごとに検証する必要がある。また、日本人英語学習者が産出したライティングプロダクトにお

ける結束性の影響を調査した先行研究はいくつかあるものの(e.g., 福田, 2020; 倉橋, 2021; 沢谷・鈴木, 2016), 特定の結束性マーカーに焦点を当てた調査であり, 英文自動解析ツールを用いた結束性指標の包括的な検証は未開拓である。以上のことから, 日本人英語学習者の産出したエッセイにおける結束性指標の影響を調査することで, 先行研究のギャップの解消に取り組む必要がある。

そこで本研究では, どのような結束性指標が日本人英語学習者の作成したエッセイに付与されたスコアの予測に寄与するかを明らかにすることを目的とした。なお, ここでのエッセイスコアとは, テキスト一貫性の判断と全体的なエッセイの質のことを指す。結束性指標の算出では, 幅広く詳細な結束性指標の算出が可能な TAACO 2.0 (Crossley et al., 2019) を用いた。本研究の分析結果から, エッセイの一貫性や質を評価するためのルーブリック開発やフィードバックの検討に寄与する内容を提示することを目指す。また, 質の高いエッセイや, まとまりのあるテキスト作成を促す指導方法への示唆を得ることも試みる。上記の目的の達成のため, 以下の検証課題 (Research Questions: RQs) を設定した。

RQ1 どのような結束性指標が, テキストの一貫性の判断に影響を及ぼすか

RQ2 どのような結束性指標が, 全体的なエッセイの質に影響を及ぼすか

3.2 データ

本研究では, アジアの国と地域の英語学習者によるスピーチやエッセイを収録した学習者コーパスである International Corpus Network of Asian Learners of English (ICNALE; <https://language.sakura.ne.jp/icnale/>) から, 学習者の CEFR レベルに加えてエッセイスコアが付与された Edited Essay (Ishikawa, 2018a) を使用した。日本の英語教育の文脈に焦点を当てているため, Edited Essay の中でも日本の大学に通う英語学習者が作成した80本(40名)のエッセイを分析対象とした。ICNALE プロジェクトに参加した学習者全員が, 以下2つのトピックについて論述型エッセイを作成した: (1) It is important for college students to have a part-time job (PTJ), (2) Smoking should be completely banned at all the restaurants in the country (SMK)。エッセイ作成時の条件は統制されており, 1つのエッセイにつき40分以内で作成するように指示を受けている。なお, 辞書の使用は禁止されているが, スペルチェッカーの使用は許可されていた (Ishikawa, 2013)。

ICNALE では, 語彙サイズテスト (Nation & Beglar, 2007) と英語外部試験 (e.g., TOEIC, TOEFL) のスコアに基づいて, 各学習者に4つの CEFR レベル (i.e., A2, B1_1, B1_2, B2) が付与されている。したがって, 本研究では CEFR の A2 から B2 レベルの英語熟達度を有する日本人が作成したエッセイが分析対象である。なお Edited Essay では, それぞれの CEFR レベルに10名の学習者が作成したエッセイが収録されているため, 4つの CEFR レベルごとに20本のエッセイの抽出が可能である (Ishikawa, 2018a)。

ICNALE Edited Essay では, 英語母語話者で学術論文の校正経験が豊富なプロのエディター5名が ESL Composition Profile (Jacobs et al., 1981) に基づいてエッセイを評価した。評価観点の情報を表2に示す。実用性や信頼性の観点からオリジナルのルーブリックを多少改良し, 5つの観点について1~12点の間で採点した。採点の信頼性のため, 8本のエッセイを抽出して各評価者が付与したスコアを共有する小規模の評価者トレーニングを実施した後に, 評価者は担当するエッセイを独立して採点した (Ishikawa, 2018a)。ルーブリックの Organization の観点では, テキストの論理的な流れに基づいて評価されるため, 一貫性の判断と類似している。したがって, Organization の観点のスコアをテキスト一貫性の判断として分析した。また, 5つの観点の合計スコアを全体的なエッセイの質として分析を進めた。

■表2: ICNALE Edited Essayで使用されたルーブリックの概要 (Ishikawa, 2018b)

観点	説明
① Content	知識が豊富で、実質的で、徹底的に練り上げられ、トピックに関連している
② Organization	十分な裏付けに基づいて、明確且つ流暢にアイデアを説明している エッセイは論理的な流れで構成されている
③ Vocabulary	語彙は洗練され、効果的に選択され、適切な範囲で使用されている
④ Language Use	複雑な構文を含み、文法的に正しい文を使用している
⑤ Mechanics	スペル、句読点、大文字、段落分けの正しい規則に基づいている

日本人英語学習者が作成したエッセイの基本情報として、Organizationスコア(テキスト一貫性の判断)とTotalスコア(エッセイの質)の記述統計を表3に示す。各スコアの記述統計をトピックごとに算出したところ、両スコア共にトピックAよりもトピックBの平均点が低いことが確認された。この結果から、トピックの要因の影響が多少表れていると言える。

■表3: 分析対象のエッセイスコアの記述統計

	エッセイ数	<i>M</i>	<i>SD</i>	Min	Max
Organizationスコア	80本	7.513 [7.210, 7.815]	1.359	4	11
トピックA (PTJ)	40本	7.950 [7.469, 8.430]	1.501	6	11
トピックB (SMK)	40本	7.075 [6.740, 7.409]	1.047	4	9
Totalスコア	80本	37.212 [36.015, 38.410]	5.383	21	52
トピックA (PTJ)	40本	39.225 [37.835, 41.066]	5.757	30	52
トピックB (SMK)	40本	35.200 [33.872, 36.527]	4.152	21	43

3.3 分析に使用する指標

エッセイの結束性指標を算出するため、英文自動解析ツールであるTAACO 2.0 (Crossley et al., 2019)を使用した。TAACOは局所的結束性(隣接文間のつながり)、大局的結束性(段落間のつながり)、テキスト全体の結束性における幅広く詳細な150以上の指標を報告する点が、既存のツールとは異なる点である (Crossley et al., 2016a)。TAACOでは表4に示す4つの結束性の特徴を報告する。なお本研究では、複数段落で構成されるエッセイを分析対象としていないため、局所的結束性とテキスト全体の結束性の指標のみ算出する。以下では、4つの結束性の特徴について説明する。各結束性指標の詳細はCrossley et al. (2016a, 2019)を参照されたい。

■表4: TAACOで報告される結束性の特徴

特徴	説明	例
語彙的重複	各品詞のレマのタイプ数を測定	The sun was bright. The day was sunny .
意味的重複	同義語の重複度を測定	The animal was huge. It was a dog .
連結詞	接続詞や語彙的従属詞などの使用を測定	Firstly , she was happy and excited.
既知性	代名詞や指示詞の出現率を測定	The girl was satisfied with what she had.

注. Crossley et al. (2016a)とKim (2022)に基づき作成。結束性の特徴に該当する英単語は太字で表示。

(1) 語彙的重複

隣接する2つの文、もしくは隣接する3つの文におけるレマの重複を測定する。テキストの総レマ数や総文数で重複したレマ数を割った値や、隣接文間で共通して使用されているレマがあるかを2値で判断した値によって、この結束性指標は数値化される (Crossley et al., 2019)。内容語と機能語のレマに加えて、名詞、動詞、形容詞、副詞、代名詞の品詞ごとに報告される。

(2) 意味的重複

Corpus of Contemporary American English (COCA) に基づいて、テキスト間の名詞と動詞の同義語の重複を測定する。例えば jump の同義語には、leap, bound, spring などの関連する単語が含まれる (Crossley et al., 2016a)。意味的重複には次の3つのモデルが含まれる: (a) テキスト内の潜在的な意味のつながりを推定する Latent Semantic Analysis (LSA), (b) テキスト内の潜在的なトピックを推定する Latent Dirichlet Allocation (LDA), (c) テキスト内の語句の意味的類似性を推定する Word2vec (Crossley et al., 2019)。

(3) 連結詞

主に2つの側面に基づいて、分析の対象となる連結詞のリストが作成されている: (a) 否定的つながり (e.g., although, unless) と肯定的つながり (e.g., actually, next) に関連する語句, (b) Halliday and Hasan (1976) の提案した文法的結束性の中でも、接続 (conjunction) に分類される結束性マーカー (Crossley et al., 2016a)。TAACO ではこれらの連結詞の出現率が測定される。

(4) 既知性

1人称代名詞 (e.g., I, us), 2人称代名詞 (e.g., you), 3人称代名詞 (e.g., he, them), 主語代名詞 (e.g., I, you, she), 数量代名詞 (e.g., many) など、さまざまなタイプの代名詞と名詞の比率が報告される。また、定冠詞 (e.g., the) や指示詞 (e.g., this, those) の出現率も算出される (Crossley et al., 2016a)。

3.4 分析方法

日本人英語学習者の作成したエッセイに付与されたスコア (i.e., Organization スコア, Total スコア) を予測するため、本研究では決定木分析と重回帰分析を実施した。分析には R ver. 4.2.2 (R Core Team, 2022) を用いた。分析手順は以下の通りである。

1. ICNALE Edited Essay (Ishikawa, 2018a) から日本人英語学習者の作成したエッセイ80本の抽出
2. TAACO 2.0 (Crossley et al., 2019) によって各エッセイの結束性指標の算出
3. 結束性指標を説明変数、スコアを目的変数とする決定木分析の実施
4. 重要度が高いと判断された結束性指標を説明変数、スコアを目的変数とする重回帰分析

TAACO では150以上の結束性指標が算出されるため、指標を選定するために決定木分析を実施した。決定木分析では rpart パッケージ (Therneau & Atkinson, 2022) を使用し、最も精度の高いノード数で決定木を作成した。なお、決定木を複数作成することで精度を高めるランダムフォレストも実施したものの、予測精度は2%程度と信頼できる結果が得られなかった。

続いて、決定木分析によって重要度が高い変数として抽出された結束性指標を説明変数に設定してエッセイスコアを予測するため、ステップワイズ法による重回帰分析を実施した。ステップワイズ法は近年批判されている手法ではあるものの (e.g., Maie et al., 2023)、結束性がライティング評価に与える影響は十分に明らかになっていない (e.g., Crossley, 2020)。したがって、本研究は探索的な検証の性質をもっていると言えるため、この手法を採用した (Eguchi & Kyle, 2020)。

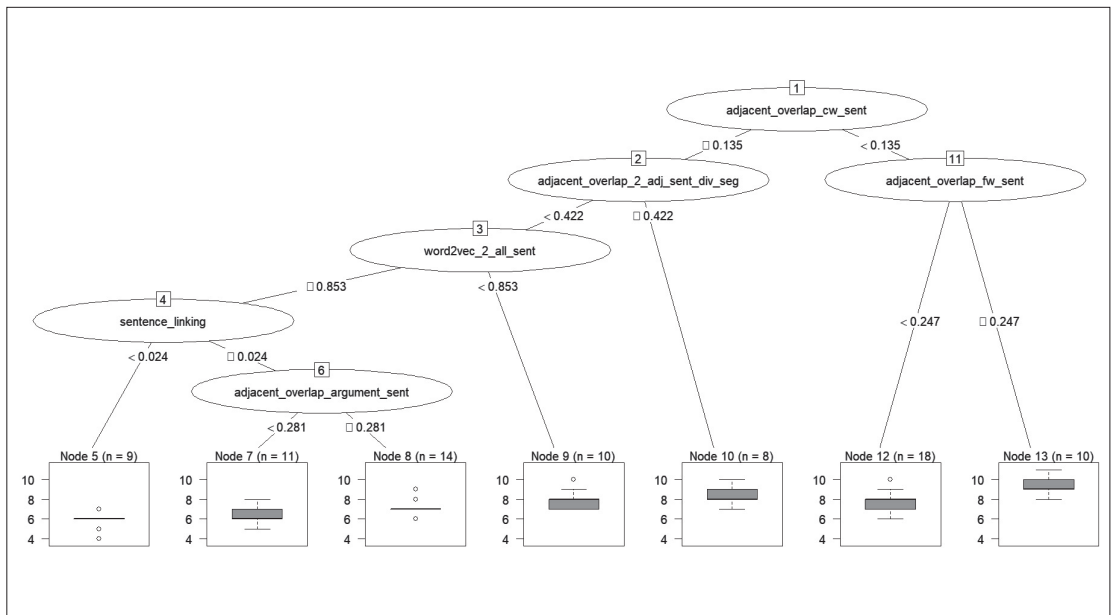
本研究で用いたデータと分析コードの詳細は資料1 (補足データ <https://osf.io/h5vpr>) を参照されたい。分析コードは平井他 (2022) や小林他 (2020) に基づいて作成した。

4 結果と考察

4.1 テキスト一貫性の判断(Organizationスコア)を予測する結束性指標

テキスト一貫性の判断を予測する決定木を図1, 分析の結果抽出された結束性指標の概要を表5に示す。決定木分析の予測精度は56.884%で, 語彙的重複の指標が重要度の高い変数として多く抽出された。

表5に示す6つの指標間の相関係数は.80以下で許容度の逆数を示すVIF (Variance Inflation Factor)の値も10以下であったため(資料1と2参照), 多重共線性の問題は確認されなかった(平井他, 2022)。重回帰分析の結果を表6に示す。2つの指標が削除され, 残った4つの指標でスコアを約17%予測することが確認された。以下では, 有意差が確認された2つの結束性指標に焦点を当てて, 考察を行う。



■ 図1: テキスト一貫性の判断(Organizationスコア)を予測する決定木

■ 表5: テキスト一貫性の判断を予測する決定木で抽出された結束性指標の概要

	結束性指標	結束性の特徴	説明
1	adjacent sentence overlap content lemmas	語彙的重複	隣接文間で少なくとも1回出現する内容語のレマタイプの数
2	adjacent sentence overlap function lemmas	語彙的重複	隣接文間で少なくとも1回出現する機能語のレマタイプの数
3	adjacent two-sentence overlap adjective lemmas (sentence normed)	語彙的重複	次の2文に少なくとも1回出現する形容詞のレマタイプの数
4	word2vec similarity (two adjacent sentences)	意味的重複	すべての隣接文(2文スパン)間の意味的類似度スコアの平均
5	sentence linking	連結詞	文を連結する機能をもつ単語の数 EX) nonetheless, therefore
6	adjacent sentence overlap noun and pronoun lemmas	語彙的重複	隣接文間で少なくとも1回出現する名詞と代名詞のレマタイプの数

■表6: 結束性指標によるテキスト一貫性の判断(Organizationスコア)を予測する重回帰分析の結果

結束性指標	<i>B</i>	95% CI	<i>SE B</i>	β	<i>t</i>	<i>p</i>
adjacent sentence overlap content lemmas	-11.665	[-18.481, -4.747]	3.374	-0.586	-3.457	.001
adjacent two-sentence overlap adjective lemmas	1.954	[0.195, 3.450]	0.742	0.271	2.633	.010
sentence linking	17.601	[-5.111, 38.769]	10.868	0.162	1.629	.108
adjacent sentence overlap noun and pronoun lemmas	5.720	[-1.694, 11.884]	2.934	0.288	1.950	.055

注. $R^2 = .208$, adjusted $R^2 = .166$.

表6から、隣接文間で内容語が多く重複しているエッセイほど、テキスト一貫性が低いと判断される傾向にあることが示された($t = -3.457$, $p = .001$)。このような特徴を含むエッセイの抜粋を以下に示す(Excerpt 1, Excerpt 2)。隣接文間で繰り返し使用している内容語レマは、太字で表記している。

高いスコアが付与されたエッセイ(Excerpt 1)では、隣接文間で重複している内容語レマがほとんどないことが確認できる。また、低いスコアが付与されているエッセイ(Excerpt 2)では、プロンプト記載の単語(part-time job)を繰り返し使用している傾向にあることも示された。以下ではトピックA(PTJ)のみを抜粋しているが、トピックB(SMK)についても同様の傾向が確認された。

Excerpt 1: 内容語レマの重複の指標が低くてスコアが高いエッセイの一部

- I agree that it is important for college students to have a part-time job. I have three reasons to support my stance. Firstly, a **part-time job** is an important chance to experience the society. Because a college is the place to prepare before going to the real society, **part-time job** is very necessary to know

(ID = W_JPN_PTJO_015_B1_2, スコア = 10, トピック = PTJ)

Excerpt 2: 内容語レマの重複の指標が高くてスコアが低いエッセイの一部

- I agree with the statement because **part-time job** gave me lots of experiences. For example, I understood following things through **part-time job**. It is hard to **make** money because I **have** to go to **part-time job** when I promised day and time, even if I don't want it. I **have** to keep time because if I delayed, I interrupted colleague's **jobs** and **make** bad relationships. I can't stay I'm center of the world because guests choice restaurant which they want to go, so I need to **make** them comfortable for they choose my shop. I **learned** from part-time **jobs** how important money and to **make** money, to **make** relationship. But I **learned**....

(ID = W_JPN_PTJO_024_B1_1, スコア = 6, トピック = PTJ)

また、内容語でも形容詞が隣接文間で重複しているエッセイでは、テキスト一貫性が高いと判断される傾向にあることが示された($t = 2.633$, $p = .010$)。このような特徴を含むエッセイの抜粋を以下に示す(Excerpt 3)。隣接文間で繰り返し使用している形容詞レマは太字で表記している。Excerpt 3では、主張を支える理由部分でトピックについての利点や欠点を説明するために形容詞(e.g., important, good, better, less)を使用する傾向にあることが確認された。このような特徴は両トピックのエッセイで見られた。

しかし、形容詞が隣接文間で重複していてもテキスト一貫性が低いと判断されるエッセイ(Excerpt 4)も確認された。このようなエッセイでは、他の形容詞(e.g., college)を繰り返し使用していたり、利点や欠点の説明で形容詞を使用していても同じ単語(e.g., bad)を繰り返し使用していたりする特徴が見られた。これは、内容語レマを繰り返し使用すると、テキスト一貫性の判断にマイナスの影響を与えるという上記の結果と一致するものである。

Excerpt 3: 形容詞レマの重複の指標が高くてスコアが高いエッセイの一部

- I think that it is **important** for college student to have a part-time job. This is because that they can learn a lot of things through their part-time jobs. For example, they work in a store, they learn that to smile is very **important**, and And this is also because that to do a lot of part-time jobs when they were college students will become a **good** memory in the future. They will enjoy their lives **better** by

(ID = W_JPN_PTJ0_022_B1_1, スコア = 10, トピック = PTJ)

- I think that smoking should be As a result of this, in the case where a restaurant does not ban smoking completely, it gathers **less** customers than if it banned smoking. This leads to the **less** benefit of

(ID = W_JPN_SMK0_009_B2_0, スコア = 8, トピック = SMK)

Excerpt 4: 形容詞レマの重複の指標が高いのにスコアが低いエッセイの一部

- I agree with this statement The second reason is that the wage for the **college** students is cheaper than other workers. So if **college** students work part time, many companies and shops do without paying higher wage to employments. But now the condition of business is very bad in Japan. Some people think because of the **college** students part-time, many unemployed people miss their opportunities to work and the unemployed rate is high. Certainly the population of workforce doesn't contain the number of **college** students and the part-time by **college** students causes ...

(ID = W_JPN_PTJ0_019_B1_2, スコア = 6, トピック = PTJ)

- I agree the idea Smoking in restaurants has two **bad** things. One is that smell of smoking is **bad** for Smell can make foods good, or otherwise it can make foods **bad**. And the other is that smoke is **bad** for health. In restaurants, there are a lot of people, babies, children, teens, adults, and olds. Smoking has **bad** influence in them all. Smoking is **bad** for health of all people, especially **bad** for

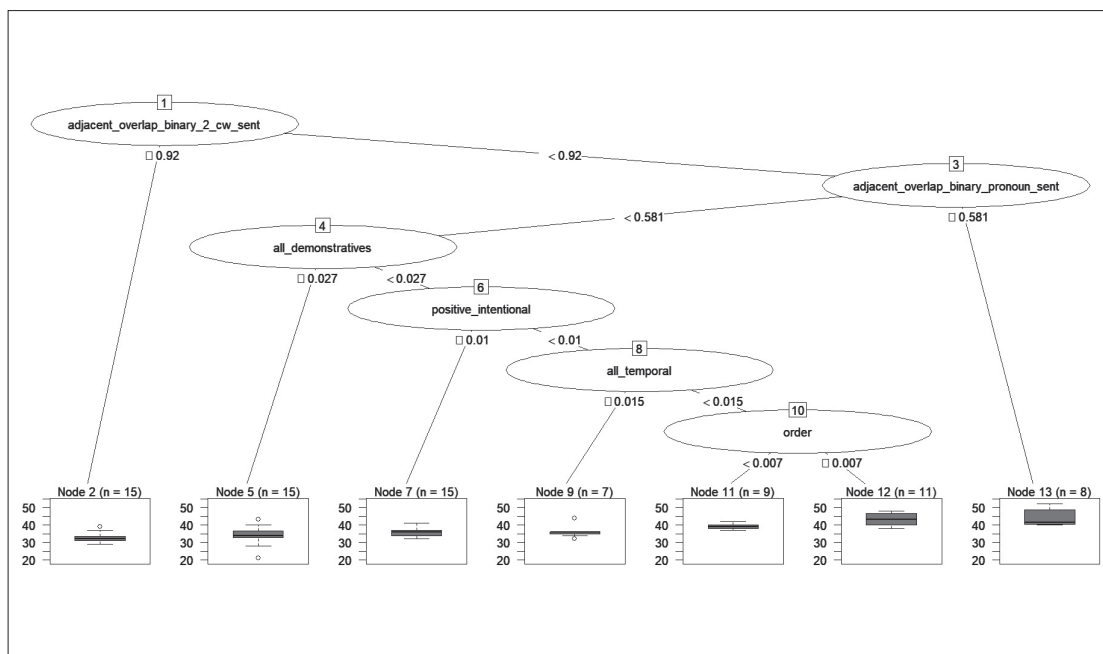
(ID = W_JPN_SMK0_018_B1_1, スコア = 6, トピック = SMK)

形容詞レマの重複の結果について、利点や欠点を明示的に表現して主張と理由の間のつながりを明確にさせることによって、テキスト一貫性の判断に影響を与えた可能性が示唆された。しかし、形容詞を使用せずとも、トピック内容の利点や欠点を説明することは容易である。そのため先行研究では、形容詞レマの繰り返し使用がエッセイの一貫性に与える影響について報告しているものは見当たらない。また、今回のデータにおいて高いスコアが付与されたエッセイで確認されることの多かった形容詞(e.g., good, bad)は、どれも洗練性の低い語彙であった。一般的に、語彙の洗練性はライティング熟達度を予測する因子として考えられているため(e.g., Crossley, 2020; Guo et al., 2013; Kojima & Kaneta, 2022), 語彙研究の知見とは矛盾する結果である。以上のことから、形容詞レマの重複の観点のみではこの結果の説明は難しく、主張を支える理由の内容や質など、他の要素が影響している可能性もあり得るだろう。実際に、今回使用した ICNALE Edited Essay での採点では、Organization は5つの観点の合計スコアと最も強い正の相関があり($r = .889$), アイデアの明確性や流暢性などの多様な要素が含まれていることが示唆されている(Ishikawa, 2018b)。つまり、テキストのつながりやまとまり以外の要素も Organization の評価に関連していた可能性がある。

4.2 エッセイの質(Totalスコア)を予測する結束性指標

エッセイの質を予測する決定木を図2、分析の結果抽出された結束性指標の概要を表7に示す。決定木分析の予測精度は57.100%で、連結詞の指標が多く抽出されたが、重要度の高い変数(決定木の上のノードにある変数)は語彙的重複の指標であった。

表7に示す6つの指標間の相関係数は.80以下でVIFの値も10以下であったため(資料3参照)、多重共線性の問題は確認されなかった(平井他, 2022)。重回帰分析の結果を表8に示す。3つの指標が削除され、残った3つの指標でスコアを約18%予測することが確認された。 $(t = -2.105, p = .039)$



■図2: エッセイの質(Totalスコア)を予測する決定木

■表7: エッセイの質を予測する決定木で抽出された結束性指標の概要

結束性指標		結束性の特徴	説明
1	binary adjacent two-sentence overlap content lemmas	語彙的重複	次の2文に何らかの内容語レマが重複する文の数
2	binary adjacent sentence overlap pronoun lemmas	語彙的重複	隣接文間で何らかの代名詞レマが重複する文の数
3	demonstratives	連結詞	指示詞の使用頻度 EX) this, that, these
4	positive intentional connectives	連結詞	目的・願望・手段などの肯定的なつながり示す 単語の数 EX) by, so, want
5	temporal connectives	連結詞	時間経過や順序を示す単語の数 EX) a consequence of, after, again
6	order	連結詞	順序を示す機能をもつ単語の数 EX) to begin with, next, first

■表8: 結束性指標によるエッセイの質(Totalスコア)を予測する重回帰分析の結果

結束性指標	<i>B</i>	95% CI	<i>SE B</i>	β	<i>t</i>	<i>p</i>
binary adjacent two-sentence overlap content lemmas	-8.012	[-15.555, 0.180]	3.806	-0.209	-2.105	.039
binary adjacent sentence overlap pronoun lemmas	7.847	[1.638, 12.531]	2.565	0.283	3.059	.003
positive intentional connectives	-182.983	[-324.696, -15.090]	74.792	-0.237	-2.447	.017

注. $R^2 = .214$, adjusted $R^2 = .183$.

以下では、有意性が確認された3つの結束性指標について考察を行う。

まず、隣接文間で内容語が多く重複しているエッセイほど、質が低いと判断される傾向にあることが示された($t = -2.105$, $p = .039$)。つまり、質の高いエッセイでは同じ内容語レマを隣接文間で繰り返し使用しない特徴が見られる。この結果は、テキスト一貫性の判断を予測するRQ1の結果や、L2ライティングにおける結束性の影響を調査した多くの先行研究(e.g., Crossley & McNamara, 2012)の結果とも一致する。したがって、内容語を繰り返し使用して局所的結束性を高めているエッセイは、テキスト一貫性やエッセイの質の判断の両方において、マイナスの影響を及ぼすことが明らかになった。

次に、内容語でも代名詞が隣接文間で重複しているエッセイでは、質が高いと判断される傾向にあることが示された($t = 3.059$, $p = .003$)。このような特徴を含むエッセイの抜粋を以下に示す(Excerpt 5)。隣接文間で繰り返し使用している代名詞レマは太字で表記している。

高いスコアが付与されたエッセイ(Excerpt 5)では、3人称代名詞theyのレマを使用して、college studentsやpeopleの繰り返し使用を避けながら理由を説明している特徴が見られた。但し、この特徴はトピックA(PTJ)のみでトピックB(SMK)では確認されなかった。一方で、代名詞が隣接文間で重複していてもエッセイの質が低いと判断されるエッセイ(Excerpt 6)も確認された。このようなエッセイでは、3人称代名詞theyのレマではなく、1人称代名詞Iのレマを過剰使用している特徴が見られた。この傾向は両トピックにおいて発見された。これは日本人英語学習者が英語母語話者と比較してIを多用することを報告している先行研究(e.g., Natsukari, 2012)と一致する結果であり、エッセイの質が低いほどこの特徴は顕著であることが示唆された。

Excerpt 5: 代名詞レマの重複の指標が高くてスコアが高いエッセイの一部

- I think that it is important for college student to have a part-time job. This is because that **they** can learn a lot of things through their part-time jobs. For example, **they** work in a store, **they** learn that to smile is very important, and that there are a lot of types of people in the society. This is also because that **their** experiences which were got by **their** part-time jobs are very useful, when **they** look for what **they** want to be in the future. **They** can do various types of part-time jobs, so it helps for **them** to look for what **they** want to be in the future. And this is because that **they** cannot do a part-time job when **they** start to work. When **they** are college students, **they** have a lot of time to do a part-time job. But when **they** start to work, **they** don't have. And this is also because that to do a lot of part-time jobs when **they** were college students will become a good memory in the future. **They** will enjoy **their** lives better by

(ID = W_JPN_PTJ0_022_B1_1, スコア = 48, トピック = PTJ)

Excerpt 6:代名詞レマの重複の指標が高いのにスコアが低いエッセイの一部

• I agree with the statement because part-time job gave **me** lots of experiences. For example, **I** understood following things through part-time jobs. It is hard to make money because **I** have to go to part-time job when **I** promised day and time, even if I don't want it. **I** have to keep time because if **I** delayed, I interrupted colleague's jobs and make bad relationships. **I** can't stay I'm center of the world because guests choice restaurant which they want to go, so **I** need to make them comfortable for they choose my shop. **I** learned from part-time jobs how important money and to make money, to make relationship. But **I** learned it is makes **me** tired to have job at the same time. **I** need time to have a rest for recovery. So **it** is interrupt **my** study to....

(ID = W_JPN_PTJ0_024_B1_1, スコア = 31, トピック = PTJ)

最後に、連結詞の結束性指標である positive intentional connectives について、肯定的つながりを示す単語数が多いほどエッセイの質が低いと判断されることが示された ($t = -2.447, p = .017$)。このような特徴を含むエッセイの抜粋を以下に示す (Excerpt 7, Excerpt 8)。positive intentional connectives の機能を持つ単語は太字で表記している。なお、この指標に該当する主な語句は以下の通りである: by, in order, make, purpose, so, want (Crossley et al., 2019)。

低いスコアが付与されたエッセイ (Excerpt 7) では、make, want, so を繰り返し使用している特徴が見られた。一方で positive intentional connectives に該当する単語を多く使用していても、質が低いと判断されていないエッセイ (Excerpt 8) も確認された。このようなエッセイでは、上記の3つの単語以外にも in order や by も使用している特徴が見られた。この傾向は両トピックで確認された。したがって、質の高いエッセイを作成するには want や make などの洗練性の低い動詞の過剰使用を抑え、so 以外の接続語を使用する点に留意する必要があることが示唆された。また、この結果は内容語レマを繰り返し使用するとテキスト一貫性やエッセイの質が低くなるという上記の結果を支持するものでもあると言える。

Excerpt 7: positive intentional connectives に該当する単語数が多くてスコアが低いエッセイの一部

• I agree with the statement It is hard to **make** money because I have to go to part-time job when I promised day and time, even if I don't **want** it. I have to keep time because if I delayed, I interrupted colleague's jobs and **make** bad relationships. I can't stay I'm center of the world because guests choice restaurant which they **want** to go, **so** I need to **make** them comfortable for they choose my shop. I learned from part-time jobs how important money and to **make** money, to **make** relationship. But I learned it is **makes** me tired to have job at the same time. I need time to have a rest for recovery. **So** it is

(ID = W_JPN_PTJ0_024_B1_1, スコア = 31, トピック = PTJ)

Excerpt 8: positive intentional connectives に該当する単語数が多いがスコアが低いエッセイの一部

• I agree with Smoking from cigarette **makes** me uncomfortable, especially when I have **So**, Japanese government should **make** smoking banned at all restaurants **in order** to save all citizens' health. From my personal experience, I believe that smoking should be banned in all restaurants in Japan. One of my friends could quite to smoke **by** the number of This **makes** her smoking less cigarettes per a day. Then she doesn't **want** to smoke anymore now. Therefore since I heard this story, I've believed that policy of smoking being banned in all restaurants **makes** it

(ID = W_JPN_SMK0_014_B1_1, スコア = 40, トピック = SMK)

5 結論

5.1 結果のまとめと教育的示唆

本研究では、TAACO 2.0(Crossley et al., 2019)によって算出した指標を用いて、日本人英語学習者のエッセイスコア(i.e., テキスト一貫性の判断を示す Organization スコア, エッセイの全体的な質を示す Total スコア)を予測する結束性指標を検証した。RQ1(どのような結束性指標が、テキストの一貫性の判断に影響を及ぼすか)では、内容語の重複はテキスト一貫性の判断にマイナスの影響を与えることが示された。また、形容詞の重複はプラスの影響を与える結果が得られたものの、これは ICNALE Edited Essay で使用されたルーブリックにおける Organization の観点の記述子に影響されている可能性が考えられる(Ishikawa, 2018a)。但し、同じ形容詞を繰り返し使用している場合には、スコアにマイナスの影響を与える傾向も見られた。

RQ2(どのような結束性指標が、全体的なエッセイの質に影響を及ぼすか)についても、内容語の重複が多いほど低いスコアが付与されていた。また、代名詞の重複もエッセイの質の予測に寄与することが明らかになったが、1人称代名詞の多用は避け、名詞の繰り返し使用を防ぐために3人称代名詞を用いることで、エッセイの質が高くなる傾向が示された。さらに、連結詞の指標である positive intentional connectives の中でも make, want, so の多用を避けることで、エッセイの質にマイナスの影響を与えないことが示唆された。つまり、同じ単語を繰り返し使用している場合には、エッセイの質が下がる傾向にあることが示唆された。

以上のことから、本研究から得られた主な結果は以下の通りである:(1)テキスト一貫性と質が高いと判断されるエッセイでは、内容語の繰り返し使用が少なく;(2)質が高いと判断されるエッセイでは、1人称代名詞の繰り返し使用が少なく、名詞の重複を避けるために3人称代名詞を使用している傾向にある。但し、代名詞の使用はトピックに影響を受ける可能性がある点に留意する必要がある。

本研究の結果に基づく、CEFR の A2 レベル以上の日本人英語学習者が作成したエッセイライティングにおける結束性に焦点を当てた指導と評価に対する示唆として、以下の2点が挙げられる。第1に、テキスト一貫性やエッセイの質を高める指導や評価では共通して、隣接文間での内容語の繰り返し使用を避けるように学習者に注意を向けさせることが望ましい。そのため、ルーブリックでは内容語の重複に関する評価項目を設けることで、同じ内容語レマの過剰使用を避けるように促す必要があるだろう。

第2に、エッセイの質を高める指導や評価では、内容語の重複に加えて代名詞の使用にも着目することが推奨される。エッセイの全体的な質の向上に焦点を当てたライティングタスクでは、1人称代名詞の過剰使用を避けて、3人称代名詞を使用して名詞の繰り返し使用を防ぐための明示的な指導や、フィードバックを提示することが望ましいだろう。

本研究では、連結詞の中でもテキスト一貫性の判断やエッセイの質の予測に、because や therefore などの談話標識の使用は影響を与えないという結果が得られた。これは、ほとんどのエッセイで談話標識が確認されたためだと考えられる。しかし、CEFR の A2 レベル以上の日本人英語学習者のエッセイであっても談話標識の出現が確認されているだけで、正確に使用できていない場合も散見された。そのため、代名詞だけではなく、談話標識の適切な使用についての指導や評価も重要事項であると言える。

5.2 本研究の限界点と今後の課題

今後は、主に以下の3点の限界点を調査することで、結束性の観点からのライティング評価や、今回の分析で得られた結果の一般化につなげていく必要がある。第1に、本研究では2つのトピックをまとめて分析を行ったため、トピックの要因を統制することができなかった。実際に2つのトピックの平均点にも

多少の開きがあることが確認されているため(表3参照), 今後はサンプルサイズを増やしてトピックごと
に調査する必要がある。これにより, 決定木分析の予測精度の向上も期待できる。

第2に, 本研究では重回帰分析にステップワイズ法を用いたことにより, 結果の解釈が難しい結束性指
標がスコアを予測する変数として抽出された場合があった。例えば, テキスト一貫性の判断を予測する
指標を検証するRQ1では, 形容詞の重複がスコアの予測に寄与することが示された。しかし, 形容詞レマ
の隣接文間での繰り返し使用がプラスの影響を及ぼすことを報告する先行研究は, 私の知る限りではな
い。結束性指標がライティング評価に与える影響は十分に明らかになっていないものの(e.g., Kojima &
Kaneta, 2022), 先行研究で示された結果の傾向や研究目的に基づいて, スコアに寄与する結束性指標を
ある程度整理した上で検証する方法も実施する必要があるだろう。

第3に, 本研究ではライティングスコアの分散の説明率が20%以下と低い結果が報告された。先行研究
においても同様の説明率であるが(e.g., Tywoniw & Crossley, 2019), これは結束性マーカーの誤用を
考慮していないことが要因の1つとして考えられる。TAACOではエッセイに出現している結束性マーカー
に基づいて指標を数値化しているため, 不適切な使用であったとしても指標値に反映されてしまう。その
ため, 結束性マーカーの誤用に焦点を当てて再度検証することも求められる。また, 結束性マーカーが使
用されるべきテキスト内の適切な位置や距離の観点からも検証することで(鈴木, 2023; 保田, 2024), ラ
イティングスコアを予測する結束性指標の特定につながるだろう。

その他の要因として, 結束性指標とルーブリックの評価方法の特性の違いも挙げられる。結束性指標
などの比率尺度は, ルーブリックによる評価尺度よりも狭い構成概念の測定に焦点を当てる(Koizumi
et al., 2022)。つまり, 個々のつながりを測定する結束性指標では評価者によるテキスト一貫性の判断
の20%の予測に止まったという結果は, 一貫性の判断には結束性以外の要素も大きく関わることを示唆
するとも考えられる。Organizationの観点では一貫性以外の要素の影響も考えられるものの(Ishikawa,
2018a), 高いスコアが付与されるエッセイでは, 洗練性の高い語彙や統語的に複雑な文を使用しながら
まとまりのあるテキストを作成することが想定される(Crossley, 2020)。そのため, 一貫したテキスト形
成には結束性以外にも関わることになる。したがって, ライティング熟達度の主要な予測因子である語彙の
洗練性や複雑性などの指標も含めた検証も今後行う必要がある。

謝辞

本研究を実施する機会をくださった公益財団法人 日本英語検定協会の皆様や選考委員の先生方に厚く
御礼申し上げます。中でも, 研究助言者である寺内一先生には非常に有益なご助言を賜りましたこと, 深
く感謝申し上げます。また, ご指導いただきました筑波大学大学院の平井明代先生と小泉利恵先生に心よ
り感謝申し上げます。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

- Bae, J., Bentler, P. M., & Lee, Y.-S. (2016). On the role of content in writing assessment. *Language Assessment Quarterly*, 13(4), 302-328. <https://doi.org/10.1080/15434303.2016.1246552>
- Cameron, C. A., Lee, K., Webster, S., Munro, K., Hunt, A. K., & Linton, M. J. (1995). Text cohesion in children's narrative writing. *Applied Psycholinguistics*, 16(3), 257-269. <https://doi.org/10.1017/s0142716400007293>
- Connor, U. (1990). Linguistic/rhetorical measures for international student persuasive writing. *Research in the Teaching of English*, 24(1), 67-87. <http://www.jstor.org/stable/40171446>
- Crossley, S. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, 11(3), 415-443. <https://doi.org/10.17239/jowr-2020.11.03.01>
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016a). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, 48(4), 1227-1237. <https://doi.org/10.3758/s13428-015-0651-7>
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016b). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1-16. <http://dx.doi.org/10.1016/j.jslw.2016.01.003>

引用文献

- Crossley, S. A., Kyle, K., & Dascalu, M. (2019). The Tool for the Automatic Analysis of Cohesion 2.0: Integrating semantic similarity and text overlap. *Behavioral Research Methods*, 51(1), 14-27. <https://doi.org/10.3758/s13428-018-1142-4>
- Crossley, S. A. & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 32(32).
- Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, 35(2), 115-135. <https://doi.org/10.1111/j.1467-9817.2010.01449.x>
- Duggleby, S. J., Tang, W., & Kuo-Newhouse, A. (2016). Does the use of connective words in written assessments predict high school students' reading and writing achievement? *Reading Psychology*, 37(4), 511-532. <https://doi.org/10.1080/02702711.2015.1066910>
- Eguchi, M., & Kyle, K. (2020). Continuing to explore the multidimensional nature of lexical sophistication: The case of oral proficiency interviews. *The Modern Language Journal*, 104(2), 381-400. <https://doi.org/10.1111/modl.12637>
- 福田陽子 (2020). 「英語学習者のライティングにおける一貫性・結束性の量的分析から質的分析へ—「読み手からみた不自然さ」を左右する要因とは—」『EIKEN BULLETIN』, 32号, 44-67. https://www.eiken.or.jp/center_for_research/pdf/bulletin/vol32/vol_32_p44-p67.pdf
- Grant, L. & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9, 123-145. [https://doi.org/10.1016/S1060-3743\(00\)00019-9](https://doi.org/10.1016/S1060-3743(00)00019-9)
- Graesser, A.C., McNamara, D.S., Louwerse, M.M. & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers*, 36, 193-202. <https://doi.org/10.3758/BF03195564>
- Guo, L., Crossley, S. A., & McNamara, D. S. (2013). Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Writing Assessment*, 18, 218-238. <https://doi.org/10.1016/j.asw.2013.05.002>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- 平井明代・岡秀亮・草薙邦広 (2022). 「教育・心理系研究のためのRによるデータ分析—論文作成への理論と実践集」東京書籍.
- Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. *Learner Corpus Studies in Asia and The World*, 1, 91-118. <https://doi.org/10.24546/81006678>
- Ishikawa, S. (2018a). The ICNALE edited essays; A dataset for analysis of L2 English learner essays based on a new integrative viewpoint. *English Corpus Studies*, 25, 117-130. https://jaecs.com/jnl/ECS25/ECS25_117-130.pdf
- Ishikawa, S. (2018b). Comparison of three kinds of alternative essay-rating methods to the ESL composition profile. *International Journal of Computer-Assisted Language Learning and Teaching*, 8(4), 32-44. <https://doi.org/10.4018/IJCALLT.2018100103>
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- Kim, J. (2022). The use of cohesive devices in Korean EFL writing across different proficiency levels. *Korean Journal of English Language and Linguistics*, 22, 1078-1100. <https://doi.org/10.15738/kjell.22.202210.1078>
- 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.
- Koizumi, R., In'nami, Y., & Jeon, E. H. (2022). L2 speaking and its internal correlates: A meta-analysis. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 proficiency: Theoretical and meta-analytic investigations* (pp. 307-158). John Benjamins. <http://doi.org/10.1075/bpa.13.10koi>
- Kojima M., Kaneta T. (2022). L2 writing and its internal correlates: A meta-analysis. In Jeon E. H., In'nami Y. (Eds.), *Understanding L2 proficiency: Theoretical and meta-analytic investigations* (pp. 109-158). John Benjamins. <https://doi.org/10.1075/bpa.13.05koj>
- 倉橋佑輔 (2021). 「英語ライティングにおける結束性」『神戸大学国際コミュニケーションセンター論集』, 17巻, 85-110. <https://doi.org/10.24546/81012587>
- Maie, R., Eguchi, M., & Uchihara, T. (2023). Arbitrary choices, arbitrary results: Three cases of multiverse analysis in L2 research. *Research Methods in Applied Linguistics*, 3(2), 100124. <https://doi.org/10.1016/j.rmal.2024.100124>
- MacArthur, C. A., Jennings, A., & Philippakos, Z. A. (2019). Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction? *Reading and Writing*, 32(6), 1553-1574. <https://doi.org/10.1007/s11145-018-9853-6>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). The linguistic features of writing quality. *Written Communication*, 27(1), 57-86. <https://doi.org/10.1177/0741088309351547>
- McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45(2), 499-515. <https://doi.org/10.3758/s13428-012-0258-1>
- McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, 22(3), 247-288. <https://doi.org/10.1080/01638539609544975>
- McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45(2), 499-515. <https://doi.org/10.3758/s13428-012-0258-1>
- Myhill, D.A. (2008). Towards a linguistic model of sentence development in writing. *Language and Education*, 22(5), 271-288.

引用文献

<https://doi.org/10.1080/09500780802152655>
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】』開隆堂.
Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
Natsukari, S. (2012). Use of I in essays by Japanese EFL learners. *JALT journal*, 34(1), 61-78. <https://doi.org/10.37546/JALTJJ35.1-4>
O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, 43(2), 121-152. <https://doi.org/10.1080/01638530709336895>
Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, 13(2), 111-129. <https://doi.org/10.1016/j.asw.2008.07.001>
Plakans, L., & Gebril, A. (2013). Using multiple texts in an integrated writing assessment: Source text use as a predictor of score. *Journal of Second Language Writing*, 22(3), 217-230. <https://doi.org/10.1016/j.jslw.2013.02.003>
Plakans, L., & Gebril, A. (2017). Exploring the relationship of organization and connection with scores in integrated writing assessment. *Assessing Writing*, 31, 98-112. <https://doi.org/10.1016/j.asw.2016.08.005>
R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
沢谷佑輔・鈴木智己 (2016).「英語ライティングにおける結束性と評価の関係性—まとまりのあるライティングを目指した実践的研究—」『北海道英語教育学会紀要』, 15巻, 35-54. https://doi.org/10.24675/helesje.15.0_35
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？—学際的な研究課題の紹介—」『日本音響学会誌』 79巻, 177-183. https://doi.org/10.20697/jasj.79.3_177
Therneau, T., & Atkinson, B. (2022). rpart: Recursive partitioning and regression trees. <https://CRAN.R-project.org/package=rpart>
Tywoniw, R., & Crossley, S. (2019). The effect of cohesive features in integrated and independent L2 writing quality and text classification. *Language Education and Assessment*, 2(3), 110-134. <https://doi.org/10.29140/lea.v2n3.151>
Tsunemoto, A., & Trofimovich, P. (2024). Coherence and comprehensibility in second language speakers' academic speaking performance. *Studies in Second Language Acquisition*, 1-23. <https://doi.org/10.1017/S0272263124000305>
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.

資料1：補足データ

本研究で用いたデータ及び分析コードは以下のリンクから入手可能です：<https://osf.io/h5vpr/>

資料2：テキスト一貫性を予測する決定木分析で抽出された結束性指標の記述統計と相関行列

結束性指標	<i>M</i>	<i>SD</i>	1	2	3	4	5	6
1. adjacent sentence overlap content lemmas	.172	.069	—					
2. adjacent sentence overlap function lemmas	.252	.076	.415	—				
3. adjacent two-sentence overlap adjective lemmas (sentence normed)	.195	.200	.254	.371	—			
4. word2vec similarity (two adjacent sentences)	.887	.029	.248	.412	.290	—		
5. sentence linking	.034	.013	-.030	.092	.147	.040	—	
6. adjacent sentence overlap noun and pronoun lemmas	.253	.077	.776	.535	.089	.353	-.089	—

資料3：エッセイの質を予測する決定木分析で抽出された結束性指標の記述統計と相関行列

結束性指標	<i>M</i>	<i>SD</i>	1	2	3	4	5	6
1. binary adjacent two-sentence overlap content lemmas	.799	.147	—					
2. binary adjacent sentence overlap pronoun lemmas	.383	.215	-.010	—				
3. demonstratives	.022	.013	.159	.174	—			
4. positive intentional connectives	.009	.008	.189	.125	-.064	—		
5. temporal connectives	.011	.007	-.099	.255	-.086	.145	—	
6. order	.008	.007	-.204	.029	-.037	-.121	.534	—

第36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

英検の長文テキストの読解に接辞の知識は どれくらい必要か？

— Morpholex Affix Profiler を用いた検討 —

研究者：駒野 樹 茨城県／筑波大学大学院 在籍

《研究助言者：齊田 智里》

概要

本研究では、英検の長文読解テキストを対象に、オンラインテキスト分析ツールのMorpholex Affix Profilerを活用して、テキスト中に含まれる派生語の割合、登場する接辞の種類や頻度、語彙カバー率への寄与度について、受験級による違いを観点に分析を行った。分析の結果、受験級が上がるにつれて、見出し語の割合は減少し、屈折語・派生語の割合が増加していることが分かった。また、接辞のレベル(Bauer & Nation, 1993)別に見てみると、高いレベルの接辞ほどテキストに含まれる割合は少なくなっていた。さらに、受験級が上がるにつれて、登場する全派生接辞のタイプ数とトークン数は増加しており、特定の派生接辞が高頻度で使用されていることが分かった。テキストの語彙カバー率への寄与度に関しては、受験級が上がるにつれて、語彙カバー率95%及び98%到達に必要な派生接辞のタイプ数は増加していた。本研究の結果は、異なる難易度のテキストとそれらに含まれる接辞の割合・種類の関係を明らかにし、語彙・読解指導やテスト作成、文章の難易度評価に示唆を与えるものである。

1 はじめに

今日の英語教育ではコミュニケーションの重要性が強調されているが、円滑なコミュニケーションを実現する上で欠かせないのが語彙知識である。小中高の外国語科においても、2017年・2018年の学習指導要領改訂により学習すべき語彙数は大幅に増加しており(文部科学省, 2017a, 2017b, 2018)、語彙を効率的かつ幅広く学ぶ必要性はますます高まっている。そのような効率的な語彙学習において有用となるのが、「接辞(affix)」の知識である。接辞の知識があれば、単語の内部構造に関する情報を活用することができるため、未知語であってもその意味を推測しやすくなり、結果的に記憶にも定着しやすくなる(磐崎, 2018)。また、未知語に遭遇した場合は、その意味を正しく推測することで記憶に定着しやすくなるという実験結果もある(Hulstijn et al., 1996)。さらに、森田(2017)は、語彙量を増やすためには派生語の知識が欠かせないとし、英語学習者に対して派生語の指導を行うべきだと主張している。

このように、語彙学習において接辞の知識は極めて重要とされるが、多種多様な接辞がある中で、それら全てを学習しようとするとは非効率的である。そのため、接辞の中でも比較的使用頻度が高く有用なものに絞って学習することが求められる。これに関連して、先行研究(e.g., Laufer & Cobb, 2020; Morita et al., 2018, 2019, 2021)では、様々なジャンルの英文テキストを対象に、どのような種類の接辞がどれほどの頻度で登場しているかを定量的に分析している。

本研究では、これらの先行研究の理論・方法論を応用して、「実用英語技能検定(以下、英検)」で出題される読解問題用の長文テキストを分析対象に、これらのテキストに含まれる接辞の特徴を明らかにする

ことを目的とした。具体的には、オンラインテキスト分析ツールのMorpholex Affix Profilerを活用して、接辞の種類や使用頻度、語彙カバー率の観点から分析を行い、異なる難易度(受験級)の間でどのような違いが見られるのかを明らかにすることで、英文読解テストの作成や文章の難易度評価、そして、英語の語彙学習や指導に関して有益な示唆を与えることを目指した。

2 先行研究の整理

2.1 語の形態論的特徴と数え方

2.1.1 語の形態論的特徴

本研究では接辞を観点にテキストを分析するため、接辞をはじめ、語の形態論的特徴についての理論を整理しておく必要がある。まず、形態素(*morpheme*)の一つとして、語の中心である基体(*base*)に付加して品詞や意味を変える働きを持つ接辞(*affix*)がある。接辞の一種である派生接辞(*derivational affix*)は、ある単語から別の単語を形成する働きを持っており、現在存在する英単語の多くは派生によって形成されたものである。例えば、形容詞*slow*に接尾辞*-ly*が付加されることで、*slowly*という副詞に派生する。西川(2006)によれば、英語には接頭辞が約155項目、接尾辞が約133項目存在するとされるが、実際には接辞と判断する基準が曖昧だったり、研究者や辞書によって判断が異なったりするとされている。先述の通り、語彙学習において単語の内部構造に着目することの効果が主張されており(e.g., 磐崎, 2018; Hulstijn et al., 1996)、このような派生接辞に関する知識は語彙学習に効果的に働くものだと考えられている。

また、屈折(*inflection*)の働きを担う屈折接辞(*inflectional affix*)は、派生接辞とは異なり、基体となる語の文法機能を変えることはない。具体例としては、動詞の過去形を示す*-ed*、三人称単数現在形の*-s*、形容詞の比較級・最上級を示す*-er*、*-est*などが挙げられる。屈折は、前後の文脈や統語的な位置関係に応じて生じるため(白畑ほか, 2019)、基本的には各屈折語が独立した見出し語として辞書に記載されることはない(不規則に変化する動詞や形容詞はこの限りではない)。対する派生語は、基本的にそれぞれ別項目として辞書に記載される(西川, 2013)。

2.1.2 語の数え方

英語によるコミュニケーションにおいて語彙知識は重要な基盤となるが、語彙知識には多様な側面が存在する。したがって、語彙知識のどの側面に着目しているかに注意しながら、学習者の語彙を測定することが求められる。門田(2014)及びNation(2022)を踏まえれば、一般的なのは量的な側面から語彙の広さ(語彙総量)を測定することであり、その場合の単語の数え方には、①延べ語数方式(*tokens*, *running words*, トークン)、②異語方式(*types*, タイプ)、③見出し語方式(*lemmas*, レマ)、④ワードファミリー方式(*word family*)がある。①は、全ての単語を出現する度に数え上げる方式である。②は、各単語を初出の際にのみ数え上げ、2回目以降は数え上げない方式である。③は、各単語の基本形に加えて、屈折形と短縮形をまとめて1単語として数え上げる方式である。④は、各単語の基本形、屈折系、短縮形とともに、その単語から派生して作られた派生形もまとめて1語として数え上げる方式である。なお、次節以降で述べる、本研究と関連の深いBauer and Nation(1993)及びLaufer and Cobb(2020)ではワードファミリー方式が用いられている。

2.2 The Affix Levels

冒頭に述べた通り、本研究ではMorpholex Affix Profilerと呼ばれるツールを用いて接辞を観点としたテキスト分析を行うが、その開発の理論的基盤となっているのがBauer and Nation(1993)が提唱した

The Affix Levelsである。The Affix Levelsとは、英語の接辞に関して、その頻度 (*frequency*) や規則性 (*regularity*)、生産性 (*productivity*) など8つの基準を総合的に踏まえて、学習・習得難易度順に7段階のレベルに分類したものである。つまり、低いレベルに属する接辞ほど学習や習得が比較的容易である一方、高いレベルに属する接辞は学習や習得が比較的困難であるということになる。なお、本レベルはリーディングの際の語彙理解に重きを置いているため、綴り字やその規則性など、書き言葉に関連する要素(基準)がより重視されている。また、存在する全ての英語の接辞が分類されているわけではない点も注意が必要である。以下、Bauer and Nationによる解説を踏まえて、8つの基準を簡潔に記す。

The Affix Levelsの作成に用いられた8つの基準

基準1: 頻度 (*frequency*)

当該の接辞が含まれている既存の単語数。The Affix Levelsで低いレベルに属する接辞ほど、より多くの単語に付加して派生語を形成していると言える。

基準2: 生産性 (*productivity*)

当該の接辞が付加されることで、新たな語が形成される可能性を表す。The Affix Levelsで低いレベルに属する接辞ほど、生産性が高く、より多くの新しい語を生み出し得ると言える。

基準3: 予測可能性 (*predictability*)

当該の接辞に含まれる意味がどれだけ予測可能かを表す。The Affix Levelsで低いレベルに属する接辞ほど、その接辞が持つ意味はより限定的、あるいは一般的であるため、接辞付加によって形成される語の意味も予測しやすいものとなる。

基準4: 基体の綴り字に関する規則性 (*regularity of the written form of the base*)

当該の接辞が付加することで生じる基体の綴りの変化が、どれだけ予測可能なものかを表す。The Affix Levelsで低いレベルに属する接辞ほど基体の綴りに影響を与えないため、その接辞を除いて残る部分は、元の基体と同じ綴りとなる。

基準5: 基体の音韻に関する規則性 (*regularity of the spoken form of the base*)

当該の接辞が付加することで、基体の音韻形式がどれだけ変化するかを表す。The Affix Levelsで低いレベルに属する接辞ほど基体の音韻形式に影響を与えないため、その接辞を除いて残る部分は、元の基体と同じ音韻形式となる。

基準6: 接辞の綴り字に関する規則性 (*regularity of the spelling of the affix*)

当該の接辞が付加することで生じる接辞の綴りの変化が、どれだけ予測可能なものかを表す。The Affix Levelsで低いレベルに属する接辞ほど、綴りが規則的でバリエーションが少ないと言える。

基準7: 接辞の音韻に関する規則性 (*regularity of the spoken form of the affix*)

当該の接辞が付加することで生じる接辞の音韻形式の変化が、どれだけ予測可能なものかを表す。The Affix Levelsで低いレベルに属する接辞ほど発音を予測しやすいため、認識が比較的容易であると言える。

基準8: 機能の規則性 (*regularity of function*)

当該の接辞が特定の品詞 (*part of speech*) の語と結合する程度、及び結合の結果として特定の品詞の語を形成する程度を表す。

これら8つの基準をもとにしてThe Affix Levelsは作成された。7段階のレベルに含まれる具体的な接辞などについては、後出の「表1. MorphoLex Affix Profilerが分析対象とする接辞のレベル別リスト」を参照されたい。また、各レベルの特徴の詳細は駒野(2024)を参照されたい。

2.3 接辞を観点とした英文テキスト分析

2.3.1 Morita et al. の一連の研究

ここでは、本研究に特に関連の深い研究として、接辞の観点から英文テキストの分析を行ったMorita et al. (2018, 2019, 2021)の一連の研究、及びLaufer and Cobb (2020)を詳細にレビューする。

Morita et al. (2018, 2019, 2021)では、日本の中学校・高等学校外国語科検定教科書を分析対象として、The Affix Levelsに分類される派生接辞の使用頻度を調査した。Morita et al. (2018)では、2016年～2019年採用の中学校英語教科書18冊(6社より3学年分ずつ)から成る教科書コーパス(総語数222,599語)を利用して、その中に登場するLevel 3～6の派生接辞のタイプ数及びトークン数を調査した。分析の結果、対象の教科書に登場する派生接辞のタイプ数、トークン数はともに少ない傾向にあり、とりわけ接頭辞(e.g., *un-*, *in-*, *inter-*, *re-*, *pre-*)は接尾辞(e.g., *-er*, *-ly*, *-y*, *-ful*, *-ity*)よりも登場頻度が低いことが示された。これを受けてMorita et al. (2018)は、教科書本文だけでは接辞の知識を習得する上で十分なインプットが得られない可能性があるため、教科書以外の教材を用いたり、接辞に関する明示的な指導を行ったりする必要があると主張している。

続くMorita et al. (2019)は、Morita et al. (2018)と同様の教科書コーパスを分析対象としつつ、抽出する接辞の種類を広げ、The Affix LevelsのLevel 7(表1参照)やそれ以外の接辞、接辞の異形態(e.g., 接頭辞 *in-* は、付加される基体に応じて *im-*, *il-*, *ir-* に変化しうる)も含めた。分析の結果、教科書の対象学年が上昇するにつれて、登場する派生接辞のタイプ数及びトークン数は増加傾向にあった。しかしながら、新たな接辞や接辞の異形態を追加したもの、全体的に登場する派生接辞のタイプ数及びトークン数は依然として少ない傾向にあった。ゆえにMorita et al. (2019)は、教科書からの接辞のインプットだけでは、学習者が自律的に接辞を学習するには不十分であることを改めて主張すると同時に、教科書を活用して接辞に関する指導を行うことを提案している。

Morita et al. (2021)では、高等学校検定教科書を対象にして、Morita et al. (2019)と同様の分析を行った。分析対象となったのは、English Communication I が12冊、同科目Ⅱが14冊、同科目Ⅲが11冊の計37冊から構成された教科書コーパスである(総語数204,158語)。分析の結果、Morita et al. (2019)の中学校教科書に登場した派生接辞は、この研究(Morita et al., 2021)が対象とした高等学校教科書においても全て登場していた。また、いずれの派生接辞もタイプ数・トークン数ともに学年の上昇に伴って増加していた。加えて、中学校段階から高校段階への移行で新たに6種類の接尾辞(*-ally*, *-atory*, *-most*, *-ory*, *-(ate+)* *able*, *-ancy*)が確認された。この研究(Morita et al., 2021)によって得られた接辞の使用頻度に関する情報は、教育現場で接辞を指導するときの指針になると述べられている。

2.3.2 Laufer and Cobb

学習者の習得語彙を測る際にワードファミリー方式を用いることについては、学習者の習得語彙を過大評価する可能性があることが指摘されている。前述の通り、ワードファミリー方式では、各単語の基本形、屈折形、短縮形、派生形も全て一括りに1語として数える。そのため、この方式による習得語彙の計測は、各単語の基本形を理解していれば、同一ワードファミリー内のその他の単語も理解しているという前提に基づいている。しかしながら、単語の基本形を理解しているからといってその派生語などを理解しているとは限らず、また、ワードファミリーの構成要素である派生語が英文テキストに占める割合が高いことを踏まえると、派生語を理解できないことで当該テキストの読解に大きく影響する可能性がある。このような問題点が、上記の指摘へとつながっている。

このような指摘を背景に、Laufer and Cobb (2020) は、オンライン分析ツール Morpholex Affix Profiler を活用して、4ジャンルの英文テキスト(①学術論文、②ニュース記事、③英語母語話者向けの名作小説、④段階別読み物教材)に含まれる見出し語(*base words*)、屈折語(*inflected words*)、派生語(*derivational words*)の頻度やテキストにおける割合、テキストの語彙カバー率95%及び98%到達への寄与度を調査した。分析に用いられた Morpholex Affix Profiler は、入力テキストから、The Affix Levels 上の接辞を含む単語を全て抽出し、その頻度(トークン数)と当該テキストに占める割合を算出する。なお Laufer and Cobb では、The Affix Levels の Level 1(見出し語)、Level 2(屈折接辞)、Level 3~6(派生接辞)が分析対象とされた(表1参照)。

分析の結果、対象テキストには平均して、Level 1(見出し語)が81.65%、Level 2(屈折接辞)が12.52%、Level 3~6(派生語)が5.60%含まれており、全体的な傾向として、英文テキストに占める派生語の割合はそれほど高くはないことが示された。また、対象テキストには特定の派生接辞(e.g., *-ly*, *-ion*, *-er*, *re-*, *un-*, *-able*)が高頻度で登場しており、それら派生接辞から形成された派生語によって語彙カバー率95%及び98%に到達しうることが示された。これらを踏まえて Laufer and Cobb (2020) は、英文テキストの読解において派生語の知識はそれほど多くは必要なく、見出し語・屈折接辞・高頻度の派生接辞の知識を備えておけば十分であると主張している。

2.4 本研究の目的と RQs

以上のように、先行研究では接辞の観点から英文テキストの分析がなされており、語彙学習や英文テキストの読解指導に対して有益な示唆が示されていた。しかしながら、これらの研究では、分析対象が英語検定教科書(Morita et al., 2018, 2019, 2021)や英語圏での読み物素材(Laufer & Cobb, 2020)となっていた。そこで本研究では、英文読解テストに用いられるテキストを対象とし、これらの研究と同様の理論的・方法論的アプローチに基づいて接辞を観点とした分析を行うことで、英文読解テストの開発や作成に関して示唆を得ることはできないかと考えた。筆者は、この点を駒野(2024)において大学入学共通テストのリーディング問題を対象に検討したが、この研究では対象とした年数が3年分(令和3年度~5年度試験)と限られていたため、結果的に分析できる問題(文章)の数が少なくなり、それゆえ文章の難易度を要因とした分析に十分な文章の数、及び語彙数を確保することができなかった。そのため、英検のように異なる難易度のテストで用いられている英文テキストを分析対象とし、難易度(受験級)による接辞の特徴を明らかにすることができれば、英文読解テストの開発や評価、語彙学習・指導に対して、英文の難易度と接辞の関係性の観点から、さらに有益な示唆が与えられるのではないかと考えた。

よって、本研究では、異なる受験級の英検の長文読解テキストを対象に、Morpholex Affix Profiler を活用して、派生語の割合、登場する接辞の種類と頻度、そして語彙カバー率の観点から分析を行うことを試みた。本研究のRQsとして、以下4つを設定した。

RQ1

異なる受験級の英検読解テキストには、見出し語、屈折語、派生語がどのような割合で含まれているか。

RQ2

異なる受験級の英検読解テキストには、The Affix Levels(Bauer & Nation, 1993)の各レベルに属する接辞がどのように分布しているか。

RQ3

異なる受験級の英検読解テキストには、具体的にどのような派生接辞が高頻度で登場しているか。

RQ4

異なる受験級の英検読解テキストにおいて、語彙カバー率95%及び98%の到達に必要な派生接辞は何か。

3 分析方法

3.1 分析対象

対象は、接辞の登場が一定数見込まれる難易度及び文章の語数を鑑みて、英検3級～1級で用いられる文章とした。これらの級について、2014年度～2023年度の10年間分の試験におけるリーディングパートの長文問題を対象とした。分析する試験の合計は、各級30回分となった。

分析対象とする問題形式は、長文の内容理解を問う「長文の内容一致選択」及び「長文の語句空所補充」とした。「短文の語句空所補充」及び「会話文の空所補充」は、読解よりも語彙・文法知識を測定する性質が強いと判断し、対象から除外した。具体的には、3級は大問4(2014年度～2016年度)及び大問3(2017年度～2023年度)、準2級は大問4・5(2014年度～2016年度)及び大問3・4(2017年度～2023年度)、2級は大問3・4(2014年度～2015年度)及び大問2・3(2016年度～2023年度)、準1級・1級は大問2・3を対象とした。なお、語彙力パー率を扱う本研究の性質を鑑み、純粋な読解テキストのみを分析対象とするため、各問における設問及び解答の選択肢は対象から除外した。また、英検で出題される問題文には、Eメール(手紙文)、説明文、評論文など複数の種類があるが、本研究ではそのようなテキストタイプは区分せずに分析を行った。上記の基準で選定した読解テキストの総語数は、225,210語(3級 19,664語、準2級 27,940語、2級 45,676語、準1級 53,656語、1級 78,274語)となった。

3.2 Morpholex Affix Profilerによる分析

接辞の分析には、Laufer and Cobb(2020)のMorpholex Affix Profiler(<https://www.lexutor.ca/cgi-bin/morpho/lex/>)を用いた。本節では、Morpholex Affix Profilerによる英文テキストの分析について、以下に簡潔に記す。より詳細な手順や基準は、Laufer and Cobb(2020)及び駒野(2024)を参照されたい。

Morpholex Affix Profilerでは、分析対象のテキストを入力すると、入力したテキストから対象の接辞が抽出され、レベル別に分類される。なお、本ツールが分析する接辞は、The Affix Levels(Bauer & Nation, 1993)に修正を加えたリスト(表1参照)の内Level 1～6に属するものである。Level 7の接辞に関しては、Laufer and Cobb(2020)に倣い、分析対象外としている。

■表1: Morpholex Affix Profilerが分析対象とする接辞のレベル別リスト(駒野, 2024, p. 34を一部修正)

接辞のLevel	各Levelの説明, または該当する接辞(太字立体は説明, 斜体は接辞, 括弧内は接辞付き語の例)
Level 1	Base words (見出し語)
Level 2	Base words + inflections ('lemmas') (見出し語 + 屈折語尾) -s (on noun or verb), -ed/-ing (on verb), -er (<i>er2</i>) / -est (on adjective), -th (on number), -en (<i>en2</i> , on irregular verb) *-er2 and -en2 are named to separate them from verb+er in Level 3 (driver) and adjective+en (widen) in Level 5
Level 3	Frequent and regular affixes with minimal change to the base word in speech or writing (高頻度かつ音韻及び綴り字の規則性が高い派生接辞) -able/ible, -er/or (on verb), -ish, -less, -ly, -ness, -th, -y, non-, un-

Level 4	Frequent orthographically regular affixes which often impose pronunciation change (admire → admirAtion) (高頻度かつ綴り字の規則性が高い派生接辞(しばしば音韻に変化が生じる)) <i>-al</i> (autumnal), <i>-ation</i> (admiration), <i>-ess</i> (fortress), <i>-ful</i> (plentiful), <i>-ism</i> (dogmatism), <i>-ist</i> (semanticist), <i>-ity</i> (solemnity), <i>-ize</i> (serialize), <i>-ment</i> (armament), <i>-ous</i> (fibrous), <i>in-</i>, <i>im-</i>
Level 5	Less frequent but regular affixes (頻度が低い, 規則性の高い派生接辞) <i>-age</i> (leakage), <i>-al</i> (arrival), <i>-ally</i> (idiotically), <i>-an</i> (American), <i>-ance</i> (clearance), <i>-ant</i> (consultant), <i>-ary</i> (revolutionary), <i>-atory</i> (confirmatory), <i>-dom</i> (kingdom; officialdom), <i>-eer</i> (black marketeer), <i>-en</i> (wooden), <i>-en</i> (widen), <i>-ence</i> (emergence), <i>-ent</i> (absorbent), <i>-ery</i> (bakery; trickery), <i>-ese</i> (Japanese; officialese), <i>-esque</i> (picturesque), <i>-ette</i> (usherette; roomette), <i>-hood</i> (childhood), <i>-i</i> (Israeli), <i>-ian</i> (phonetician; Johnsonian), <i>-ite</i> (Paisleyite, also chemical meaning), <i>-let</i> (coverlet), <i>-ling</i> (duckling), <i>-ly</i> (leisurely), <i>-most</i> (topmost), <i>-ory</i> (contradictory), <i>-ship</i> (studentship), <i>-ward</i> (homeward), <i>-ways</i> (crossways), <i>-wise</i> (endwise; discussion-wise), <i>ante-</i> (anteroom), <i>anti-</i> (anti-inflation), <i>arch-</i> (archbishop), <i>bi-</i> (biplane), <i>circum-</i> (circumnavigate), <i>counter-</i> (counter-attack), <i>en-</i> (encage; enslave), <i>ex-</i> (ex-president), <i>fore-</i> (forename), <i>hyper-</i> (hyperactive), <i>inter-</i> (inter-African; interweave), <i>mid-</i> (midweek), <i>mis-</i> (misfit), <i>neo-</i> (neo-colonialism), <i>post-</i> (post-date), <i>pro-</i> (pro-British), <i>semi-</i> (semi-automatic), <i>sub-</i> (subclassify; subterranean), <i>un-</i> (untie; unburden).
Level 6	Frequent but irregular affixes (often with significant change to base word) (頻度は高いが, 規則性が低い派生接辞(しばしば基体に変化が生じる)) <i>-able</i> (inscrutable), <i>-ee</i> (lessee), <i>-ic</i> (spastic), <i>-ify</i> (mollify), <i>-ion</i> (superstition), <i>-ist</i> (solipsist), <i>-ition</i> (transition), <i>-ive</i> (restive), <i>-th</i> (breadth), <i>-y</i> (calumny), <i>pre-</i>, <i>re-</i>.
Level 7	Classical roots and affixes (古典語由来の語根と接辞) <i>-ar</i> (circular), <i>-ate</i> (electorate), <i>-et</i> (packet, casket), <i>-some</i> (troublesome), <i>-ure</i> (departure, exposure), <i>ab-</i> (abnormal), <i>ad-</i> (admixture), <i>com-</i> (commiserate), <i>de-</i> (demist), <i>dis-</i> (disintegrate), <i>ex-</i> (out ~ external), <i>in-</i> (in ~ internal), <i>ob-</i> (obsequious), <i>per-</i> (perspective), <i>pro-</i> (in front of ~ precede), <i>trans-</i> (transmogrification)

注. Morpholex Affix Profiler のウェブサイトに掲載されたものをもとに筆者作成。Level 7 も記載しているが、本研究の分析対象ではないことに留意されたい。

次いで、Morpholex Affix Profiler による実際の分析例を簡潔に示す。今回の分析例では、2023年度第1回英検1級の長文読解テキストを入力した。テキストを入力したのち分析を開始すると、当該テキストの分析結果が出力される(図1)。

MorphoLex Affix Profiler v.5

NEW - OUTPUT SUMMARY AGGREGATED IN DATA FILE (request from button at output bottom)

For 2023年度第1回点検 1級 (3,868 words $\geq$ 3,885 w/ contractions split)

Morpholex found 664 affixed types / 1042 tokens in input of 3885 for 24.87% of text by tokens

[illegible]

図1: Morpholex Affix Profilerの分析結果の出力画面

Reasoning behind the by-affix profile

There are all 63 different derivational suffixes in 2023年度第1回英検_1級 (3885 words), sorted most to least:

61_~ly 29_~er 28_~al 25_~ment 20_re~ 19_~ity 16_~ation 15_in~ 11_un~ 11_~ive 10_~y 9_dis~ 8_~able 8_~ist 8_~ial 8_~ization 7_~ic 6_~ful 6_~age 6_~ance 6_en~ 4_~ness 4_~ized 4_~a~
4_~ent 3_~ists 3_~ous 3_~en 3_~ship 3_~ition 3_~ical 2_~or 2_~ities 2_~ize 2_~ism 2_~ence 2_~acy 2_~ways 2_~ant 2_pro~ 2_mis~ 1_~less 1_~istic 1_~ability 1_lm~ 1_~ce 1_~ry
1_~hood 1_~atic 1_~eer 1_~ary 1_~ite 1_~ward 1_bi~ 1_mid~ 1_~ically 1_~ative 1_~itive 1_~etic 1_~atory 1_de~ 1_pre~

The text is 3885 words, 95% of the text is $3885 \times .95 = 3683$ words, and 98% of the text is $3885 \times .98 = 3799$ words

Base words are 2959 and inflections are 580 so there are 3539 non-derived words

Derived words needed to reach or exceed 95% of lexis are 3683 - 3539 = 144

Derived words needed to reach or exceed 98% of lexis are 3799 - 3539 = 260

Base + inflected words + 5 deriv. affixes are 95%+ of text:

61_~ly 29_~er 28_~al 25_~ment 20_re~

Base + inflected words + 8 more deriv. affixes are 98%+ of text:

19_~ity 16_~ation 15_in~ 14_un~ 11_~ive 10_~y 9_dis~ 8_~able

■ 図2: MorphoLex Affix Profiler による語彙力バー率の算出方法

図1の上段は、入力テキストの総トークン数、抽出された接辞付き語のタイプ数及びトークン数を示している。続く中段は左から、①「(1)BY B+N LEVEL with 95 and 98% cut-offs highlighted」、②「(2)BY CLASS」、③「(3)BY AFFIX」の3パートに分かれている。①は、The Affix Levelsの各レベルに属する見出し語、屈折接辞、派生接辞それぞれのトークン数、及び当該テキストの総トークン数に占める割合を示している。②は、見出し語、屈折接辞、派生接辞それぞれのトークン数、及び当該テキストの総トークン数に占める割合を示している。③は、当該テキストの総トークン数の95%及び98%に到達するために必要とされる派生接辞を列挙している（なお、語彙カバー率算出の詳細については、図2に示した出力情報の別パートを参照されたい）。今回の分析例では、テキスト総トークン数3,885語の内、その95%は3,683語（ 3885×0.95 ）、98%は3,799語（ 3885×0.98 ）であり、それぞれの語数に達するのに十分なトークン数を有する派生接辞が提示される仕組みになっている。最後に、下段は抽出された屈折接辞（Level 2）及び派生接辞（Level 3～6）をレベル別に示している（接尾辞は赤字、接頭辞は青字で表記される）。

RQ1（見出し語、屈折語、派生語の割合）については、図1中段の②「(2)BY CLASS」で示されるLevel 1（見出し語）、Level 2（屈折語）、Level 3～6（派生語）の割合を級ごとに参照・比較することで明らかにする。RQ2（各レベルに属する接辞の分布）は、図1中段の①「(1)BY B+N LEVEL with 95 and 98% cut-offs highlighted」で示されるThe Affix LevelsのLevel 2以上の割合について、級ごとに参照・比較することで明らかにする。RQ3（高頻度で登場する派生接辞）は、図1下段及び図2で示されるLevel 3～6（派生接辞）について、級ごとに参照・比較することで明らかにする。RQ4（語彙カバー率への寄与）は、図1中段の③「(3)BY AFFIX」及び図2の出力情報について、級ごとに参照・比較することで明らかにする。

4 分析結果と考察

4.1 見出し語、屈折語、派生語の割合（RQ1）

RQ1への対応として、分析対象とした各受験級のテキストにおいて、見出し語、屈折語、派生語がそれぞれどれくらいの割合で含まれているのかを調査した。表2は、その結果をまとめたものである。一般的な傾向として、見出し語や屈折語と比較して形態論的な複雑性が高い派生語は、テキストの難易度が上昇するほどテキストに含まれる割合が増加すると予想された。この予想と一致して、今回対象とした英検読解テキストにおいても、3級から1級へと受験級が上がるにつれて、見出し語の割合は減少している一方、派生語の割合は増加していることが明らかになった。派生語の割合に着目してみると、2級から準1級では2.84ポイント増加しており、その差が比較的大きいことが分かる。また、見出し語の割合にも着目すると、同様に2級から準1級では4.63ポイント減少しており、その差が比較的大きい。これらの結果や先行研究での知見を踏まえると、形態論的な観点から述べれば、英検の読解テキストは2級から準1級にかけてその難易度が大きく上昇しているとも言える。

■表2: 各受験級のテキストにおける見出し語・屈折語・派生語の割合 (%)

受験級 (総語数)	見出し語	屈折語	派生語
3級 (19,664語)	88.54	8.87	2.26
準2級 (27,940語)	83.32	13.47	3.06
2級 (45,676語)	80.44	15.29	5.22
準1級 (53,656語)	75.81	17.74	8.06
1級 (78,274語)	75.09	16.81	9.68

屈折語に関しても、派生語と同様、全体的には受験級の上昇に伴ってその割合が増加傾向にあることが明らかになった。これは、受験級が上昇するにつれて、動詞の活用形(過去形、過去分詞)や分詞(分詞による後置修飾や分詞構文)など、動詞の屈折形がより多く用いられることを考えれば、妥当な結果であると言える。ただし、準1級から1級にかけては、屈折語の割合がやや減少していた。これは、難易度の高い文章においては、より難易度の高い派生語の割合が高くなる傾向にあるため、屈折語の割合が頭打ちになったと考えることができるかもしれない。

本研究の結果を Laufer and Cobb(2020)の結果と比較すると、Laufer and Cobbで分析対象とされた4ジャンルテキストの内、形態論的に最も複雑であった学術論文では、見出し語が75.34%、屈折語が16.88%、派生語が7.78%だった。本研究で対象とした英検読解テキストでは、準1級で見出し語が75.81%、屈折語が17.74%、派生語が8.06%、1級で見出し語が75.09%、屈折語が16.81%、派生語が9.68%であり、Laufer and Cobbにおける学術論文よりも形態論的な複雑性がやや高く、難易度の高いテキストであると判断できるだろう。また、大学入学共通テストのリーディング問題を対象とした駒野(2024)の結果と比較すると、共通テストでは見出し語が81.51%、屈折語が12.60%、派生語が5.46%であり、英検2級がこれと最も近い値であった。ただし、大学入学共通テストは全ての大問の英文を含めた場合の割合であることに留意されたい。

4.2 各レベルに属する接辞の分布(RQ2)

RQ2の対応として、分析対象とした各受験級のテキストにおいて、The Affix Levelsの各レベルに属する接辞がどれくらいの割合で分布しているのかを調査した。表3は、その結果をまとめたものである。The Affix Levelsでは、高いレベルに属する接辞ほど習得が難しく、そうした習得の難しい接辞がテキストに占める割合は少ないとされる。つまり、レベルの上昇に伴い、各レベルの接辞がテキストに含まれる割合は減少していく(Level 1から徐々に割合が減少していく)と考えられる。本研究で対象とした英検読解テキストにおいても、各級について、レベルの上昇に伴って各レベルに属する接辞のテキストに含まれる割合が減少する傾向が見られた。これは、Laufer and Cobb(2020)や駒野(2024)とも同じ傾向である。

■表3: 各受験級のテキストに登場する接辞付き語のレベル別割合 (%)

受験級(総語数)	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
3級(19,664語)	88.54	8.87	1.04	0.78	0.34	0.10
準2級(27,940語)	83.82	13.47	1.40	0.89	0.42	0.35
2級(45,676語)	80.44	15.29	1.99	1.66	0.74	0.84
準1級(53,656語)	75.81	17.74	2.91	2.58	1.34	1.23
1級(78,274語)	75.09	16.81	3.00	3.21	1.77	1.70

また、派生接辞に当たる Level 3～6に着目してみると、受験級の上昇に伴い、いずれのレベルでも当該レベルに属する接辞の割合は上昇していた。高頻度かつ規則性が高い Level 3の派生接辞は、1級を除く受験級においても、派生接辞の中で最も割合が高かった。そのため、1級を除く受験級においては、Level 3の派生接辞の知識が読解上で重要であると考えられる。また、Level 3の派生接辞は、Laufer and Cobb(2020)で最も割合が高かった学術論文で2.50%だったことを踏まえると、とりわけ準1級と1級においてその割合が高かった。

一方で、受験級が上がるにつれて、高いレベルの派生接辞の割合が Level 3と比べて相対的に増えていることも示された。つまり、英検テキストの読解においては、受験級が上がるほど、派生接辞の中でも高

いレベルに属するものに関する知識を要すると言えるだろう。また、読解テキストの難易度について検討する際、当該テキストに含まれる派生語の割合を調整するとともに、派生語の中でもより高いレベルに属する接辞の割合を相対的に増減させることも有用になる可能性がある。

4.3 高頻度で登場する派生接辞(RQ3)

RQ3の対応として、分析対象とした各受験級のテキストにおいて、登場する全ての派生接辞を抽出し、タイプ数とトークン数を集計した。表4は、各受験級のテキストに高頻度で登場する派生接辞の内、頻度順10位までを抜粋して示したものである。なお、登場した派生接辞のタイプ数及びトークン数の総計は、3級では32タイプ、444トークン、準2級では54タイプ、855トークン、2級では77タイプ、2407トークン、準1級では97タイプ、4401トークン、1級では106タイプ、7708トークンだった。これらの結果から、受験級が上がるにつれて、登場する派生接辞の種類や頻度が大きく増加することが明らかになった。また、難易度の高い級のテキストでは、各派生接辞の頻度も比較的高いことが明らかになった。

■表4: 各受験級のテキストに高頻度で登場する派生接辞 (頻度順10位までを抜粋)

受験級	派生接辞(トークン数)
3級	-er (106), -ly (63), -al (57), -ous (41), -ful (26), -y (22), -ent (18), -ation (14), -ist (12), bi- (11)...
準2級	-er (148), -ly (141), -al (55), -y (54), -ful (44), -ation (34), -ent (32), -ous (27), -ment (27), -ist (26)...
2級	-ly (351), -er (318), re- (141), -ment (140), -ist (131), -al (130), -ists (111), dis- (72), -ent (67), -ant (63)...
準1級	-ly (700), -al (363), -er (352), -ment (236), re- (170), dis- (132), un- (127), -ive (124), -ation (123), -ity (117)...
1級	-ly (1085), -al (654), -er (446), -ment (403), -ation (289), -ity (272), re- (272), un- (216), -ive (215), in- (212)...

頻度順10位に含まれる派生接辞を具体的に見てみると、-ly, -er, -alは全ての級において含まれていた。-mentは3級を除く4つの級で、-ationは2級を除く4つの級において含まれていた。-ent, -ist, re-は3つの級で、-ous, -ful, -y, dis-, un-, -ive, -ityは2つの級において含まれていた。これらを踏まえると、特定の派生接辞が幅広い受験級のテキストにおいて高頻度で登場していたと言える。中でも、-ly, -er, -alは分析対象とした全ての級に登場していたことから、当該テキストを読解する上で特に重要な派生接辞であると言える。また、-ous, -ful, -yは3級及び準2級で頻度順10位に含まれているのに対して、un-, -ive, -ityは準1級及び1級で頻度順10位に含まれている。The Affix Levelsによれば、前者3つの派生接辞はLevel 3～4に属する一方で、後者3つの派生接辞はLevel 4～6に属する。したがって、難易度の高い級のテキストほど、The Affix Levelsにおいて上位に属する派生接辞がより多く用いられていることが分かる。これは、RQ2の結果と合わせて、The Affix Levelsの妥当性を支持する結果でもある。

次いで、先行研究の結果と比較する。表5は、Laufer and Cobb(2020)の4ジャンルテキストに登場し、語彙カバー率95%及び98%への到達に必要な派生接辞(頻度順10位まで)をまとめたものである。なお、表5は4ジャンルテキストに登場した全派生接辞の内、各語彙カバー率到達に必要なものを集計した一覧であることから、あくまでも参考として参照されたい。表6はMorita et al.(2018, 2019, 2021)の一連の研究で対象とした中学校・高校英語検定教科書に登場した派生接辞(頻度順10位まで)をまとめたものである。そして表7は、駒野(2024)の共通テスト「英語(リーディング)」に登場した派生接辞(頻度順10位11項目まで)をまとめたものである。本研究で対象とした全ての受験級に登場した3種類の派生接辞(-ly, -er, -al)は、Laufer and Cobbの4ジャンルテキスト、Morita et al.の中学校・高校教科書、及び駒野の共通テスト「英

語(リーディング)」でも共通して登場していた。このことから、派生接辞 *-ly*, *-er*, *-al* は、様々なジャンルの英文テキストを読解する際に広く有用であると言えるだろう。

■表5: Laufer and Cobb (2020) の4ジャンルテキストに登場し、語彙カバー率95%及び98%への到達に必要な派生接辞 (頻度順10位までを抜粋)

	派生接辞(トークン数)
4ジャンル テキスト	<i>-ly</i> (3132), <i>-ion</i> (826), <i>-er</i> (596), <i>-y</i> (529), <i>-al</i> (437), <i>un-</i> (387), <i>-age</i> (364), <i>-ness</i> (344), <i>re-</i> (269), <i>-ity</i> (225)

注. 駒野(2024)に掲載したものを修正。

■表6: Morita et al. (2018, 2019, 2021) の中学校・高校英語検定教科書に登場した派生接辞 (頻度順10位までを抜粋)

	派生接辞(トークン数)
中学校教科書 (Morita et al., 2018, 2019)	<i>-er</i> (1085), <i>-ly</i> (716), <i>-y</i> (484), <i>-ful</i> (427), <i>-ese</i> (423), <i>-ion</i> (379), <i>-al</i> (284), <i>-ive</i> (273), <i>-ity</i> (262), <i>-ous</i> (196)
高校教科書 (Morita et al., 2021)	<i>-ly</i> (2025), <i>-ion</i> (1207), <i>-er</i> (1161), <i>-al</i> (833), <i>-ese</i> (415), <i>-ity</i> (393), <i>-an</i> (370), <i>-ation</i> (347), <i>-y</i> (340), <i>-ment</i> (299)

注. 駒野(2024)に掲載したものを修正。

■表7: 駒野(2024)の共通テスト「英語(リーディング)」に登場した派生接辞 (頻度順10位11項目までを抜粋)

	派生接辞(トークン数)
共通テスト 「英語(リーディング)」	<i>-ly</i> (209), <i>-er</i> (179), <i>-al</i> (88), <i>-ation</i> (82), <i>re-</i> (78), <i>-y</i> (57), <i>-ment</i> (49), <i>inter-</i> (46), <i>-ful</i> (38), <i>-or</i> (37), <i>-ant</i> (37)

4.4 語彙カバー率への貢献(RQ4)

RQ4の対応として、分析対象とした各受験級のテキストにおいて、特に高頻出で、当該テキストの語彙カバー率への寄与度が高い派生接辞を調査した。表8はその結果をまとめたものである。3級では、各語彙カバー率に到達するために、派生接辞は必要でないことが明らかになった。準2級では、95%到達に必要な派生接辞は見られなかったが、98%到達には *-er*, *-ly* の2種類の派生接辞が必要であることが分かった。2級では、95%到達に必要な派生接辞が見られなかったが、98%到達には *-ly*, *-er*, *-ment*, *-ist*, *re-*, *-al* の6種類の派生接辞が必要であることが分かった。準1級では、95%到達に *-ly*, *-al*, *-er*, *-ment* の4種類が、98%到達には *-er*, *-al*, *-ment*, *re-*, *un-* など計18種類の派生接辞が必要であることが分かった。1級では、95%到達には *-ly*, *-al*, *-er*, *-ment*, *re-*, *-ation*, *-ity*, *-ist* の8種類が、98%到達には *-ation*, *-ive*, *in-*, *-ity*, *re-* など計21種類の派生接辞が必要であることが分かった。

■表8: 各受験級のテキストの語彙カバー率95%及び98%到達に必要な派生接辞

受験級	語彙カバー率95%		語彙カバー率98%	
	タイプ数	派生接辞(トークン数)	タイプ数	派生接辞(トークン数)
3級	0		0	
準2級	0		2	-er (148), -ly (56)
2級	0		6	-ly (351), -er (318), -ment (106), -ist (96), re- (91), -al (30).
準1級	4	-ly (700), -al (117), -er (70), -ment (37).	18	-er (283), -al (246), -ment (199), re- (148), un- (90), -ity (84), -ive (78), dis- (70), -ation (57), -y (46), -ness (43), -ist (40), in- (22), -ful (22), -or (15), -ities (12), -able (11), en- (10).
1級	8	-ly (1085), -al (654), -er (386), -ment (234), re- (73), -ation (49), -ity (34), -ist (30).	21	-ation (240), -ive (215), in- (196), -ity (192), re- (191), un- (190), -ment (169), -or (161), -ial (130), -ist (97), -y (91), -ic (71), dis- (66), -er (60), -ists (34), -able (32), en- (20), -ization (20), -ful (17), -ness (15), -ous (13).

上記の結果をまとめると、受験級が上がるにつれて、語彙カバー率95%及び98%到達に必要な派生接辞のタイプ数は増加していることが明らかになった。とりわけ、2級から準1級にかけては、各語彙カバー率の到達に必要な派生接辞の種類が大幅に増加していた。

5 結論と今後の課題

5.1 結論

本研究では、テキスト中に含まれる派生語の割合、登場する接辞の種類や頻度、語彙カバー率への寄与度を観点に、オンラインテキスト分析ツールのMorphoLex Affix Profilerを活用して、異なる受験級の英検の長文読解テキストを分析した。以下、各RQをもとに本研究で得られた結果をまとめる。

RQ1では、各受験級のテキストに含まれる見出し語、屈折語、派生語の割合をそれぞれ調べた。その結果、受験級が上がるにつれて、見出し語の割合は減少しているのに対し、派生語の割合は増加していることが分かった。また、屈折語についても、受験級の上昇に伴ってその割合が増加する傾向が見られた。とりわけ準1級と1級に関しては、先行研究で対象とされたテキストや、隣接する英検2級のテキストと比較して派生語の割合が特に大きくなっていることから、形態論的な複雑性が高く、読解の難易度も高くなり得るテキストであると考えられる。

RQ2では、The Affix Levelsの各レベルに属する接辞が、各受験級のテキストにどのような割合で分布しているのかを調べた。各受験級について、先行研究と同様に、The Affix Levelsの上昇に伴い、各レベルに属する接辞がテキスト中に登場する割合は減少傾向にあることが分かった。次いで、The Affix Levelsの各レベルに着目してみると、派生接辞に当たるLevel 3～6では、受験級が上がるにつれてその割合が増加していた。特にLevel 3の接辞の割合は、いずれの受験級でも最も高かったことから、テキストを読解するうえで重要な接辞のタイプであると言えるだろう。一方で、Level 5・6といった高いレベルの接辞に関して、受験級の上昇に伴ってLevel 3と比較してその割合が相対的に増えていることから、よ

り難易度の高い文章を読解する際には、高いレベルの接辞の知識も重要であると考えられる。ただし、この点については、以下で述べる本研究の限界点も参照されたい。

RQ3では、各受験級のテキストに登場する全ての派生接辞を抽出し、それぞれのタイプ数とトークン数を集計した。その結果、受験級が上がるにつれて、登場する派生接辞のタイプ数とトークン数が増えていることが分かった。また、各派生接辞のトークン数も、難易度が高い級のテキストほど多くなっていることが分かった。さらに、各級の頻度順10位以内に含まれる派生接辞に着目すると、特定の派生接辞が幅広い受験級のテキストにおいて高頻度で登場していた。特に *-ly*, *-er*, *-al* などの派生接辞は、先行研究で分析対象とされた多様なジャンルのテキストにも広く登場していたことから、英文読解において有用な派生接辞であると言えるだろう。

RQ4では、各受験級のテキストにおいて、特に高頻度で登場し、当該テキストの語彙カバー率への寄与度が高い派生接辞を調べた。その結果、受験級が上がるにつれて、語彙カバー率95%及び98%到達に必要な派生接辞のタイプ数は増加していることが分かり、特に2級から準1級にかけて大幅に増加していることが明らかになった。

本研究で得られた結果は、学習者向けに作成された異なる難易度の英文テキストと、それらに含まれる接辞の割合・種類の関係を明らかにするものである。これらの知見が、英語学習者に対する語彙指導や読解指導、そして英文読解テスト開発やテキストの難易度評価において有用となれば、筆者として幸甚である。

5.2 今後の課題

最後に、本研究の限界点と課題について述べておかななくてはならない。まず、駒野(2024)でも指摘されているように、派生接辞について学習者が理解していることは、必ずしもその派生接辞を伴う単語を完全に理解していることを意味しないということに留意が必要である。その派生語のもととなる見出し語(基体)の知識が無ければ、その単語の意味を完全に理解することはできず、この点を踏まえたうえで本研究の結果は慎重に解釈されるべきものである。

次に、テキストの難易度は、本研究で扱った接辞をはじめとする形態論的知識のみに依存しないという点に留意する必要がある。テキストの難易度を決める要素としては、その他に単語自体の難易度や構文の複雑さ、テキストが取り扱う題材が学習者にとってどれほど親しみのあるものなのかなどが考えられ、様々な観点からテキストの難易度は測られるべきである。そのため、本研究の結果は、テキストの難易度を推測するための一観点にすぎないと捉えるべきだろう。

最後に、テキストタイプやテキストのトピックごとの形態論的特徴を明らかにしていない点が挙げられる。既述の通り、英検で出題される問題文には、Eメール(手紙文)、説明文、評論文など様々なタイプがあるが、本研究ではそれらを区別せずに一括りにして分析を行った。また、テキストのトピックについても、分析の際に特段考慮していない。しかしながら、テキストタイプやテキストのトピックによって、用いられる単語などに違いが見られると推測される。例えば、テキストのトピックに関連の深い単語はそのテキスト中で多用される傾向にあるため、結果的にそうした単語に付加している接辞が偏って多く計上されることが考えられる。したがって、テキストタイプやテキストのトピックを区別して分析すれば、それぞれに異なる特徴が見られる可能性がある。これらの限界点と課題を踏まえ、接辞などの形態論的特徴をはじめ、より幅広い視点からテキスト分析が行われることが期待される。

謝辞

本研究を実施するにあたり、今回このような貴重な機会をくださいました公益財団法人 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この場をお借りして厚く御礼申し上げます。特に、研究助言者の齊田智里先生には、本研究の実施及び報告書作成に関して有益なご助言・ご指導をいただきました。心より感謝申し上げます。また、筑波大学大学院の名畑目真吾先生には、本研究の構想から実施、執筆に

至るまで、懇切丁寧なご指導をいただきました。誠にありがとうございました。そして、これまでお互いに支え合いながら研究を進めてきた同期の皆様や先輩・後輩の皆様にも、心より感謝申し上げます。最後になりますが、本研究を実施するにあたりサポートいただきました全ての皆様に深く感謝申し上げます。ありがとうございました。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

- Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253-279.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign-language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327-339.
- 磐崎弘貞 (2018).「第19章 語彙と辞書検索指導」望月昭彦 (編著).『新学習指導要領にもとづく英語科教育法 第3版』(pp. 241-253) 大修館書店.
- 門田修平 (編著) (2014).『英語のメンタルレキシコン:語彙の獲得・処理・学習』第3版松柏社.
- 駒野樹 (2024).「接辞を観点とした英文読解テキストの分析—大学入学共通テスト「英語 (リーディング)」を対象として—」筑波大学大学院教育学学位プログラム修士論文 (未公開).
- Nation, I. S. P. (2022) *Learning vocabulary in another language* (3rd Ed.). Cambridge University Press.
- Laufer, B., & Cobb, T. (2020). How much knowledge of derived words is needed for reading? *Applied Linguistics*, 41(6), 971-998.
- 文部科学省 (2017a).「【外国語活動・外国語編】小学校学習指導要領 (平成29年告示) 解説」Retrieved from https://www.mext.go.jp/content/20220614-mxt_kyoiku02-100002607_11.pdf (最終閲覧日2024.08.20).
- 文部科学省 (2017b).「【外国語編】中学校学習指導要領 (平成29年告示) 解説」Retrieved from https://www.mext.go.jp/content/20210531-mxt_kyoiku01-100002608_010.pdf (最終閲覧日2024.08.20).
- 文部科学省 (2018).「【外国語編 英語編】高等学校学習指導要領 (平成30年告示) 解説」Retrieved from https://www.mext.go.jp/content/1407073_09_1_2.pdf (最終閲覧日2024.08.20).
- 森田光宏 (2017).「接辞を用いた語彙指導の理論と実践:派生語指導のススメ」『KELES ジャーナル』第2巻, 50-54.
- Morita, M., Uchida, S., & Takahashi, Y. (2018). Prefixes and suffixes in Japanese junior high school English textbooks. *Proceedings of Asia Pacific Corpus Linguistics Conference*, 4, 335-339.
- Morita, M., Uchida, S., & Takahashi, Y. (2019). The frequency of affixes and affixed words in Japanese junior high school textbooks: A corpus study. *ARELE: Annual Review of English Language Education in Japan*, 30, 129-143.
- Morita, M., Uchida, S., & Takahashi, Y. (2021). The frequency of affixes and affixed words in Japanese senior high school English textbooks: A corpus study. *ARELE: Annual Review of English Language Education in Japan*, 32, 81-95.
- 西川盛雄 (2006).『英語接辞研究』開拓社.
- 西川盛雄 (2013).『英語接辞の魅力:語彙力を高める単語のメカニズム』開拓社.
- 白畑知彦・富田祐一・村野井仁・若林茂則 (2019).『英語教育用語辞典 第3版』大修館書店.

第36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

ディベートスキルとジャッジスキルの 向上を目指した高校生英語ディベート初心者への 効果的なフィードバック

研究者:坂口 寛子 福岡県／福岡県立春日高等学校 教諭

《研究助言者:齊田 智里》

概要

本研究は、「論理・表現」の重要な活動の一つであるディベート指導に焦点を当てた。通常、ディベートでは、試合後、ジャッジよりディベーターに口頭で判定結果と共にフィードバックが与えられるが、本研究は、口頭以外の方法によるフィードバックが、ディベートスキルやジャッジスキルの向上に効果があるかを明らかにすることをその目的とした。具体的には、英語ディベート初心者の異なる3グループにそれぞれ、教師から口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバックのいずれかのフィードバックを与え、各フィードバックが、ディベートスキルとジャッジスキル向上に効果があるか、また、教師ジャッジの判定決断までの思考過程の伝達に効果があるかを質問紙調査と自由記述調査により明らかにした。その結果、グラフィックオーガナイザー・フィードバックにはディベートスキルにおいて有意な差が見られたが、ジャッジスキルにおいては有意な差が見られたフィードバックはなかった。また、自由記述回答からは、3種類のフィードバックは、ロジックの吟味、異なる主張間の整合性の確認、議論展開の理解といった、ジャッジの視点や方法に関する学びを促すことも分かった。英語教育ではこれまで、ディベートのフィードバックのあり方についてはほとんど研究がなされてこなかったが、本研究結果より、勝敗よりもスキルや思考方法の習得が重要視される教育ディベートでは、今後、口頭フィードバック以外のフィードバック方法の授業導入の必要性が示唆された。

1 はじめに

2022年より実施されている学習指導要領において、「論理・表現」は、「話すこと（やり取り）」「話すこと（発表）」「書くこと」を中心とした発信力育成を目指しており（文部科学省、2018）、ディベートもその目標達成の重要な活動の一つである。「論理・表現Ⅰ」ではディベートを「賛成又は反対の立場を明確にして、自分の立場の意見や考えがより妥当性や優位性があることを述べる活動」（文部科学省、2018, p. 87）とし、最終的な目的を聴衆の説得としている。ディベートにおける説得すべき聴衆とはジャッジを意味し、ディベーターが、聴衆であるジャッジによる肯定・否定側の意見の妥当性や優位性の判断について理解することは、「論理・表現」の目標を達成する上で重要である。しかしながら、高校での英語ディベート指導に関する先行研究は少ないのが現状である。そこで本研究は、高校入学後初めて英語ディベートに取り組む英語ディベート初心者を対象に、2・3学期の授業において即興型ディベート指導を実施し、試合後、教師から異なる三つの各グループに、口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバックのいずれかのフィードバックを行い、各フィードバックが、ディベートスキルとジャッジ

スキル向上に効果があるか、また、教師ジャッジの勝敗の判定を下すまでの思考過程を分かりやすく伝えることができるか、を明らかにすることを目的とした。

2 研究の背景

2.1 ディベートの定義・種類・教育効果

ディベートは、「与えられた論題について肯定側と否定側に分かれて、一定のルール、進行の方法、制限時間などに従って、自分の立場の正当性を聴衆に受け入れてもらうことを目的として議論する対立型の言語活動」(文部科学省, 2018, p.113)と定義されている。チームメンバー同士が協力しながら、議論を対立させることにより自らの立場の優位性や妥当性を論理的に述べ、聴衆であるジャッジを説得することをその目的としている。ディベートの種類には、教育ディベート、競技ディベート、実社会ディベートの3種類がある(茂木, 2001)。日本の中学・高校で実施されている教育ディベートには、事前調査型英語ディベートと即興型英語ディベートの2種類があり、さらに各英語ディベートには授業英語ディベートと競技英語ディベートがある(河野, 2021)。ディベートの本来の目的はジャッジの説得であるが、教室で行われる教育ディベートでは、相手を論破し勝利することよりも、その目的を達成するまでの過程において身につけられる論証スキルや論理的思考力といった、スキルや思考方法の習得に力点が置かれる。また、ディベート指導は、言語力伸長のみならず、(1)批判的思考力(Akerman & Neale, 2011)、(2)論理的思考力(松本・鈴木・青沼, 2008)、(3)協調性(河野, 2021)、(4)幅広い知識(加藤, 2020; 河野, 2021)、(5)情報収集・活用能力(松本, 1996)、(6)迅速な思考力(松本, 1996; 松本・鈴木・青沼, 2008)などの習得・育成も期待できる。

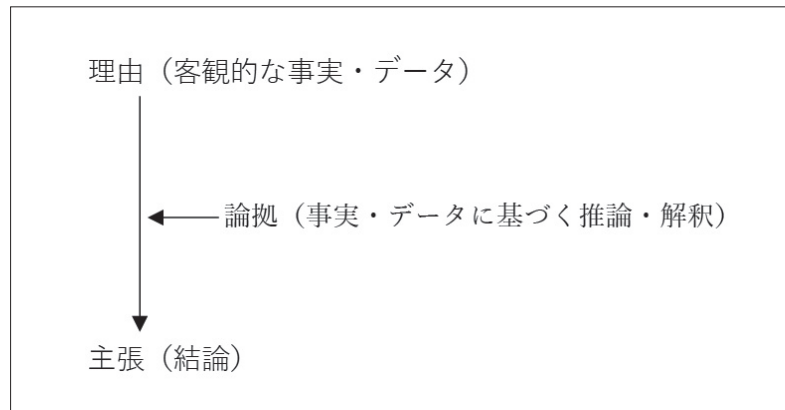
2.2 高校における英語ディベート

言語力向上に加え、実社会で必要な思考力育成等も期待できる英語ディベートであるが、残念ながら、その実施率は低い。高校入学前に英語ディベートを経験したことのある高校生は非常に少なく(坂口, 2022)、多くの高校生が高校入学後に初めて英語ディベートに取り組む。また、英語ディベートを「よく行う」「ときどき行う」と回答した高校教員は5.3%に留まり(Benesse, 2016)、教師のディベート経験不足による心理的・技術的ハードルの高さ(檜尾, 2011)、限られた授業時間やクラスサイズ(市川, 2018)などの要因が指摘されている。さらに、2021年より使用されている中学3年生英語検定教科書内のディベート指導に関して、中井(2022)は、論理的な思考方法についての指導の不十分さを指摘している。このことから、高校では、試合形式だけでなく、三角ロジックなどを用いたロジック、主張を裏づける根拠や論拠の適切さ、内容的つながりを吟味する指導を行う必要があると思われる。三角ロジックとは、理由、論拠、主張の3要素から構成されるロジックであり、鶴田(2017)¹は、理由を客観的な事実・データとし、論拠を事実・データに基づく推論・解釈とし、論拠の役割を主張と理由の橋渡しとしている(図1)。横山(2006)は、ディベートをロジカルなコミュニケーションとし、ディベートでは“*How and why?*”(どのように、なぜ)を論証する論証責任を果たす必要があるとしているが、日本人にとって当たり前ではない論証責任を果たすには、ロジックの基本を身につける必要があるとしている。また、福澤(2018)²は、論証とは主張を何らかの理由によって裏づけることとしている。

注

1 鶴田は、その著書の中で、本研究における理由を根拠、論拠を理由という言葉で表している。

2 福澤は、その著書の中で、本研究における理由を根拠という言葉で表している。



■図1: 三角ロジック(鶴田, 2017, p. 16をもとに作成)

2.3 ディベートにおけるフィードバック

ディベーターへのフィードバックの効果について、佐藤(2011)は、日本語教育におけるディベート指導の中で、クラスメイトや教師によるフィードバックは、ディベーターの改善点認識や不安軽減につながるが、ディベーターが、試合中に自身の試合展開をモニターすることは容易なことではないため、試合後の客観的な試合の振り返りの支援が必要であるとしている。また、ディベーターは試合結果だけでなく、試合中のどの議論が評価され、あるいは評価されなかったのか、判定に至るまでのジャッジの思考過程を知りたいとされており(是澤, 2002)、フィードバックがディベーターを育成する上で果たす役割は大きいと思われる。さらに教育ディベートでは、生徒自身がディベートのジャッジを行うことが多いため、生徒がジャッジの議論の見方や判定の決定方法について学ぶことが求められる。このことから、試合の振り返りや、ジャッジの思考過程の伝達という点において、フィードバックは、ディベーターとジャッジの育成において重要であると考えられる。

2.4 フローシート

フローシートとは、ジャッジとディベーターが使用する、それぞれのスピーカーが話した内容をメモするためのシートである(中川, 2019)。各列に立論、質問、反論の各ディベーターの発話内容を左列から右列に順番で記入し、矢印で関連のある内容同士をつなぐことにより、議論の流れの把握が可能となる。小林(2019)は、フローシート記入について、ジャッジは詳細情報を記入することが望ましいが、ディベーターは、その役割を果たす上で必要な情報のみを書き留めることが望ましく、また、記入上のポイントとして、賛成・反対派の発言内容を区別するために、異なる色のペンで適宜、略語や日本語を使用することを指摘している。フローシートの具体例は図2(P.66)を参照されたい。

2.5 グラフィックオーガナイザー

グラフィックオーガナイザーとは、線、円、ボックスを用い、階層、原因・結果、比較・対比、循環・直線といった方法により情報描写を行い、コミュニケーションや情報理解を促す視覚支援である(Ellis & Howard, 2007)。グラフィックオーガナイザーは、学習者の情報の分類、情報間の関連性の確認、テキスト構造の把握、情報内のメインアイデアの抽出を容易にするとされ、リーディング力やライティングスキルの向上、理科などの科目の内容理解の促進、思考の活性化にも効果があるとされているが(Praveen & Rajan, 2013)、高次思考への効果はまだ明らかにされていない(Ellis & Howard, 2007)。グラフィックオーガナイザー

の具体例は図3(P.66)を参照されたい。

3 研究の目的

本研究の目的は、高校生英語ディベート初心者異なるグループそれぞれに、教師ジャッジから口頭フィードバック、フローシートを使用したフィードバック、グラフィックオーガナイザーを使用したフィードバックの3種類の異なるフィードバックを行い、各フィードバックによって、参加者のディベートスキルとジャッジスキルが向上するか、また、教師ジャッジの判定を下すまでの思考過程を分かりやすく伝えることができるかを明らかにすることである。

RQ1 口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバックは、高校生英語ディベート初心者のディベートスキルにどのような効果をもたらすか。

RQ2 口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバックは、高校生英語ディベート初心者のジャッジスキルにどのような効果をもたらすか。

RQ3 口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバックは、高校生英語ディベート初心者教師ジャッジの思考過程を分かりやすく伝えることができるか。

4 研究の方法

4.1 参加者

参加者は、高校入学後初めて英語ディベートを経験する、普通科在籍の高校1年生96名(口頭フィードバック29名、フローシート・フィードバック34名、グラフィックオーガナイザー・フィードバック33名)である。3クラスを対象とし、各クラスに異なるフィードバックを与えた。英語力は、校外模試結果より推察する限り、英検準2級レベルの参加者が大半である。海外留学経験者や日常的に英語を使用している参加者はいない。また、ほぼ全員が4年生大学への進学を希望している。なお、参加者には、本研究の目的を説明した上で、調査結果は本研究以外の目的には使用されないこと、同意なく第三者に提供しないこと等を口頭と書面で説明し、同意を得た。

4.2 ディベート指導

英語ディベート初心者を対象とした指導においては、以下の7点に配慮した。(1)アイデア創出、発話量増加、反論思考といった思考力活性化を図り、ディベートへの心理的障壁を下げるため、ペアディベート、サークルディベートなど様々な形式のディベートを行う。(2)参加者にとって背景知識がある論題を選択する。(3)試合の流れの把握、相手チームのスピーチの聞き取りを支援するため、フローシートにサインポスト(主張内容を端的に表す言葉)を記入する枠や、主張や理由を区別して記入することができる欄を設ける。(4)参加者のレベルに応じて、試合中の準備時間の長さや頻度を柔軟に設定する。(5)Google JamboardやGoogle Documentの共有機能を用い、協働学習を促す。(6)ロジックの型の理解のため三角ロジックを用いる。(7)再反論を含まない、立論、質疑応答、反論による通常の試合より短い20分版の試合を行う。各回の授業内容(表1)、授業のトピック・論題(表2)、試合の流れと時間配分(表3)を次頁以降に示す。

■表1: 授業内容

回	授業内容	授業目標
1	1. 日本語と英語による三角ロジックの理解 2. ブレインストーミング 3. パラグラフライティング	1. 三角ロジックを理解することができる 2. 論理的なつながりに気をつけ、1段落を書くことができる
2	1. ディスコースマーカー 2. パラグラフライティング	1. パラグラフライティングに慣れることができる 2. ディスコースマーカーを使い、結束性のある文章を書くことができる
3	1. 準備時間ありのペアトーク 2. 準備時間ありのチェントーク 3. チェントークのトピックについてのライティング	1. 間違いを恐れずに英語でたくさん話することができる 2. クラスメイトと英語でやり取りをすることができる 3. 少しの準備があれば英語のスピーチをすることができる 4. 話した内容をライティングし、正確さを高めることができる
4	1. 準備時間なしのペアトーク 2. ペアディベート 3. ペアディベートの論題についてのライティング 4. 日本語即興型ディベートの動画視聴	1. 論理性を意識し、スピーキングやライティングをすることができる 2. 即興性を高めることができる 3. 話した内容をライティングし、正確さを高めることができる 4. 即興型ディベートのざっくりしたイメージを持つことができる
5	1. 日本語によるペアディベート 2. 英語によるペアディベート 3. ペアディベートの論題についてのライティング	1. 論理性を意識し、相手を説得するために英語でスピーチをすることができる 2. 話した内容をライティングし、正確さを高めることができる
6	1. ディベートについての説明 2. 質問の定型表現 3. フローシートの取り方の練習 4. 質問を考える 5. シナリオミニディベート	1. 立論・質疑応答・反論が何であるのかを理解することができる 2. 質問の定型表現を理解することができる 3. 教師による立論を聞き、フローシートを記入しながら、立論に対する質問をできるだけ多く考えることができる 4. 立論、反論のフォーマットを理解することができる 5. サインポストが何であるのかを理解することができる 6. シナリオミニディベートにより、ディベートの流れを理解することができる
7	1. 6種類の反論表現の使い分け 2. 反論理由 3. 日本語サークルディベート 4. 試合に向けた準備	1. 適切な反論表現を選択することができる 2. 反論の理由を考えることができる 3. 日本語により立論・反論・要約の練習をすることができる
8	試合に向けた準備	グループで協力して、Google Classroomの共有機能を用い、英語ディベートの準備をすることができる
9	全参加者による英語ディベート	事前に相手の立論原稿を把握し、それに対する質問や反論をグループで協力して考え、英語ディベートを行うことができる
10	試合に向けた準備	1. ALTから海外における学校の清掃事情について英語で説明を聞き、理解することができる 2. 前回の英語ディベートの反省点を踏まえ、試合の準備をすることができる
11	全参加者による英語ディベート	前回とは異なる役割で英語ディベートをすることができる
12	1. モデルディベートとジャッジの練習 2. 全参加者による英語ディベート	1. ディベーターは、ジャッジや相手チームの立場に立ち、分かりやすさに配慮したディベートを行うことができる 2. 代表2チームによる試合のジャッジに挑戦することができる
13	1. モデルディベートのフィードバック 2. 事前調査	1. 論理性の高い立論と反論を考えることができる 2. 相手の主張の弱い点を見つけようとする可以尝试
14	全参加者による英語ディベート	事前に相手の立論原稿を把握し、それに対する質問や反論をグループで協力して考え、英語ディベートを行うことができる
15	モデルディベートとジャッジの練習	1. これまで学んだことを意識し、論理性、重要性を意識した試合を行うことができる 2. 前回の学びをいかし、ジャッジをすることができる
16	1. モデルディベートのフィードバック 2. 事後調査	教師からのフィードバックを参考に、ジャッジの仕方について理解することができる

■表2: 授業のトピック・論題

回	トピック・論題
1	ライティング 日本の公立高校は、土曜日に授業をすべきか
2	ライティング All of the high school students should have a driver's license.
3	ペアトーク 1. Have you bought anything recently? 2. Which is more difficult for you, math or English? 3. Do you often LINE your friends? 4. Do you have any plans for this weekend? チェーントークとライティング 5. Do you think that all high schools should take every Friday off? 6. Do you think that Fukuoka is the best place to live in? 7. Which do you think is better, home-cooked food or fast food? 8. Which do you think is better, apartments or houses?
4	ペアトーク 1. Which country do you want to visit? 2. Do you often watch YouTube? 3. What did you do last weekend? 4. Do you want to get a driver's license at the age of 18? ペアディベートとライティング 5. All of the elementary school students should use SNS. 動画視聴 未成年のエナジードリンク購入を禁止するべきだ(灘中学・高校ディベート部)
5	ペアディベートとライティング Kasuga Senior High School should establish STARBUCKS in its school area.
6	教師による立論のみのディベート 1. Basketball is better than soccer. シナリオミニディベート 2. Japanese government should ban tobacco.
7	日本語サークルディベート 1. All of the elementary school students should use SNS. 2. Kasuga Senior High School students should be allowed to go to school by motorcycle after the age of 16.
8 9	英語ディベート Kasuga Senior High School students should be allowed to go to school by motorcycle after the age of 16.
10 11 12 13	英語ディベート In some foreign countries, schools employ cleaning staff to clean classrooms. Do you think schools in Japan should make students clean their classrooms?
14 15 16	英語ディベート Online shopping is better than shopping in stores.

■表3: 試合の流れと時間配分

Affirmative	Negative	Time (mins)
Constructive		2
	Constructive	2
Preparation		6
	Cross examination	2
Cross examination		2
Preparation		2
	Rebuttal	2
Rebuttal		2
Total		20

4.3 手順

2023年9月から3月までの半年間、高校1年生「論理・表現 I」の授業において計16回(1回50分)の指導を行った。教師より、高校入学後初めて英語ディベートを経験する高校生3グループそれぞれに、主張内容の重要性・論理性を端的に表した○△×の記号を用いた口頭フィードバック(表4)、フローシート・フィードバック(図2)、グラフィックオーガナイザー・フィードバック(図3)のいずれかが与えられた。各フィードバックは、モデルディベート後の13回目と16回目の授業において、各グループに計2回与えた。フィードバックを与えた際のモデルディベートの論題は、“In some foreign countries, schools employ cleaning staff to clean classrooms. Do you think schools in Japan should make students clean their classrooms?”と“Online shopping is better than shopping in stores.”の二つとした。フィードバックの提示方法については、ジャッジは通常、口頭でフィードバックを行い、試合内容を記入したフローシートをディベーターに提示することはない。しかし、本研究のフローシート・フィードバックでは、教師ジャッジが記入したフローシートを配布し、参加者はフローシートを見ながら教師ジャッジからのフィードバックを聞いた。また、グラフィックオーガナイザー・フィードバックについては、松下・前田・田中(2022)の対話型論証モデル(図4)を参考に、Google Jamboardで作成したグラフィックオーガナイザーをプロジェクターで教室前方のスクリーンに投影し、参加者はスクリーン上のグラフィックオーガナイザーを見ながら、教師ジャッジからのフィードバックを聞いた。グラフィックオーガナイザーの中身については、三角ロジックの各要素は異なる色の枠(主張はピンク、理由は水色、論拠は緑)で示され、松下・前田・田中(2022)の対話型論証モデルでは立論の横に対立意見(相手サイドの主張)が配置されている。しかし本研究では、反論が、立論の理由、論拠、あるいは主張と理由・論拠のつながりの3箇所の内、どこをアタックしているのかを明確にするため、立論の上に対立意見を位置づけた。

■表4: 「生徒の学校清掃」に関する口頭フィードバックの一部

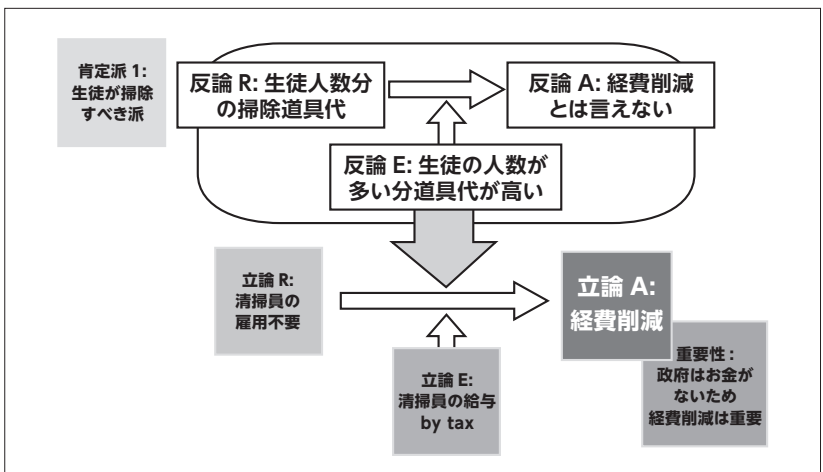
否定側(論理性)△→×・(重要性)△→×

否定側1つ目の立論「勉強時間の確保」の論理構成は、主張が「勉強時間の確保」、その理由が「掃除時間により勉強時間が減るから」というもので、この主張と理由をつなぐ論拠は説明されていませんでした。この主張は「勉強時間が減り成績が下がる」ということでその重要性が説明されていました。この立論に対し、肯定側は「生徒が清掃時間を勉強時間に充てるとは限らない」ため、「成績が低下するとは言えない」と反論しました。この反論は、主張「勉強時間の確保」と理由「掃除時間により勉強時間が減るから」をつなぐ論拠がない点をアタックしたものでした。清掃することにより勉強時間が減るとは言えないため、成績低下が起こる確率も低くなったと判断しました。(以下省略)

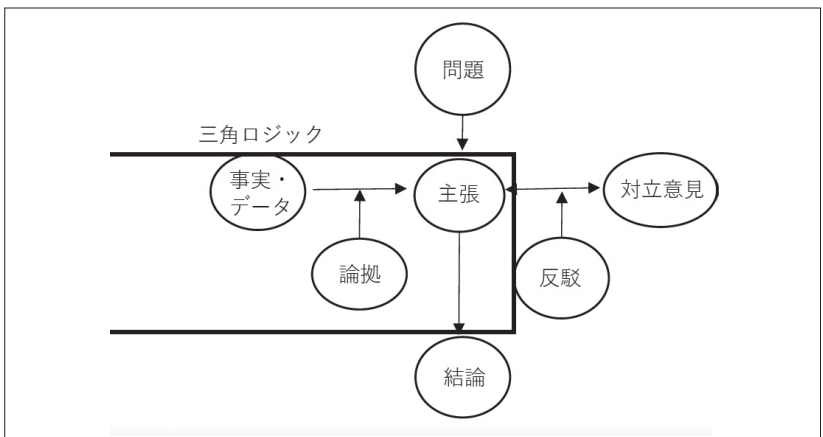
Proposition: Online shopping is better than shopping in stores A

1 Affirmative constructive speech	2 Preparation	3 Cross Examination from neg. to aff.	4 Negative constructive speech	5 Preparation
Signpost 1: 利便性 Reason 1 ①いつでも買い物が可能 ②店舗に行く必要がない "E" to support Reason 1 ①店では、仕事、勉強、家事などで多忙であるため、営業時間内に買い物できることは限らないが、オンラインショップは常に開店しているため、いつでも買い物が可能 ②店舗に足を運ぶことが難しい人(ex. 赤ちゃん連れの人)にも買い物を楽しむ公平な機会を確保できる Imp: 買い物なしでは生活できないため、オンラインショッピングは私たちの生活を保障してくれる	論理性: 〇→〇 重要性: 〇→〇 1. オンラインショッピング時に、何度も通信障害が生じたことを知っていますか? →購入に支障をきたす程の大きな通信障害はありませんでした		Signpost 1: 買い物の間違いが起きない Reason 1 店舗購入では、欲しいと思う商品を確実に入手でき、期待はずれの商品を回避できる "E" to support Reason 1 オンラインでは商品説明や写真を参考に商品を確認するが、店舗では自分の目で現物を確認できる Imp: 期待通りの商品を顧客に届けることで、顧客の期待に沿うことができる→期待を裏切ると顧客の失望を招く	論理性: 〇→〇 重要性: 〇→△

■図2: 「オンラインショッピング」に関するフローシート・フィードバックの一部



■図3: 「生徒の学校清掃」に関するグラフィックオーガナイザー・フィードバックの一部 (各四角内の枠内は要素により色分けされる)



■図4: 対話型論証モデル (松下・前田・田中, 2022, p. 4)

4.4 ディベートスキル

ディベーターは、対立型インタラクションの中で、ジャッジを論理的に説得する。その際重要なのは、論題を多角的に分析し、肯定・否定両方の立場から考え、与えられた情報を鵜呑みにせず複眼的に考える批判的思考力と、理由や論拠が伴う自身の意見を主張できるロジックである。この二つの力はディベーターのみならずジャッジにも必要な力である。多くのディベート大会、検定教科書等のジャッジシートには、ロジックと主張の伝え方(デリバリー)に関する2種類の評価の観点が含まれているが、本研究では英語ディベート初心者への指導内容の焦点化を図るため、ロジックに焦点を当てたディベートスキルを茂木(2013)、内藤・西村(2018)、鶴田(2017)をもとに作成した。以下が、本研究におけるディベートスキルである。

- (1) 相手の立論と反論中の理由や論拠に誤りがないかを考える
- (2) 相手の立論と自分たちの反論の整合性を考える
- (3) 自分たちの反論中の主張、理由、論拠のつながりが適切であることを考える
- (4) 自分たちの立論と相手の立論の相違点を考える
- (5) 相手の立論中の主張、理由、論拠のつながりが適切であるかは考えない(反転項目)
- (6) 利害関係者の立場から論題について考えてみるといったような、多角的な視点から論題について考える
- (7) 相手の立場に立ち、相手からの反論を予想する
- (8) 自分たちの立論を、主張、理由、論拠をもとに考える
- (9) 相手の立論中の主張、理由、論拠の区別を考える
- (10) 相手の立論と反論中の理由や論拠の情報が不足していないかを考える
- (11) 自分たちの反論を、主張、理由、論拠をもとに考える
- (12) 自分たちの立論を、主張、理由、論拠のつながりを意識し考える
- (13) 自分たちの立論と反論中の理由や論拠の情報が不足していないかを考える

(茂木, 2013; 内藤・西村, 2018; 鶴田, 2017をもとに作成)

4.5 ジャッジスキル

近年、授業内で生徒同士が英語ディベートの試合や審判を行う機会が増加しているため、生徒のジャッジ育成の必要性は高まっている(小林, 2021)。しかし、残念ながらジャッジの方法論、教育方法についての研究はほとんど行われていないのが現状である(久保, 2017)。ジャッジは判定を下す際、個人的な知識や経験、考え方ではなく、議論内容のみを判断材料とし(安井, 1994)、ディベーターにより提示されたメリットとデメリットが発生するか、そしてそのメリットとデメリットが、どれ程重要、あるいは深刻であるかを検証し判定を下す(内藤・西村, 2018)。ジャッジの役割について Hanson(1997)は、ジャッジには肯定派・否定派の両サイドの議論を公平に評価した上でその優劣を判断し、ディベーターに改善点を述べる役割があるとしており、是澤(2002)もジャッジの decision maker(判定者)と educator(教育者)の二つの役割を指摘している。しかし、高校生英語ディベート初心者が、この二つの役割を担うことはハードルが高い。まずは、議論の見方や捉え方を身につけることが重要であると考え、本研究におけるジャッジスキルを、藤岡・上條(1996)、内藤・西村(2018)、矢野(2020)、安井(1994)をもとに作成し、以下のように設定した。

- (1) 英語の流暢性ではなく、議論の中身をもとに判定を考える
- (2) 両者の一部分ではなく試合全体のやり取りから、判定を考える
- (3) 個人的な考えや知識を判定材料にせず、試合の議論内容から判定を考える
- (4) 立論・反論中の主張、理由、根拠のつながりが適切であることを考える
- (5) 反論が、立論に対する適切な反論になっているかを考える

- (6) 立論・反論中の主張, 理由, 論拠を区別する
- (7) 立論・反論中の理由や論拠に誤りがないかを考える
- (8) 立論・反論中の理由や論拠の情報に不足がないかを考える
- (9) 両者の立論に出されたメリット・デメリットが本当に生じるかを考える
- (10) 肯定側から出されたメリットの重要性, 否定側から出されたデメリットの深刻性を比較し, 判定を考える

(藤岡・上條, 1996; 内藤・西村, 2018; 矢野, 2020; 安井, 1994をもとに作成)

4.6 データ分析

- (1) 3種類の各フィードバックがディベートスキルとジャッジスキルに与える効果を検証するため, 4件法による質問紙調査(1:あてはまらない 2:あまりあてはまらない 3:ややあてはまる 4:あてはまる)(表5と表6)の事前・事後調査の結果に対し, HAD(清水, 2016)を用い二要因分散分析を行った。
- (2) 参加者の教師ジャッジの思考過程への気づきを検証するため, 2種類の調査を実施した。一つ目は4件法による8つの質問項目から構成される質問紙調査(1:そう思わない 2:あまりそう思わない 3:ややそう思う 4:そう思う)(表7), 二つ目は, 「教師からのフィードバックについてどう思いましたか」という質問に対しての自由記述による調査である。質問紙調査結果に対してHAD(清水, 2016)を用いて一要因分散分析を行い, 自由記述の結果をKH Coderにより分析した。

■表5: ディベートスキルに関する質問紙調査

1	相手の立論と反論中の理由(R)や論拠(E)に誤りがないかを考える	1	2	3	4
2	相手の立論と自分たちの反論の整合性を考える	1	2	3	4
3	自分たちの反論中の主張(A), 理由(R), 論拠(E)のつながりが適切であるかを考える	1	2	3	4
4	自分たちの立論と相手の立論の相違点を考える	1	2	3	4
5	相手の立論中の主張(A), 理由(R), 論拠(E)のつながりが適切であるかは考えない	1	2	3	4
6	利害関係者の立場から論題について考えてみるといったような, 多角的な視点から論題について考える	1	2	3	4
7	相手の立場に立ち, 相手からの反論を予想する	1	2	3	4
8	自分たちの立論を, 主張(A), 理由(R), 論拠(E)をもとに考える	1	2	3	4
9	相手の立論中の主張(A), 理由(R), 論拠(E)の区別を考える	1	2	3	4
10	相手の立論と反論中の理由(R)や論拠(E)の情報に不足していないかを考える	1	2	3	4
11	自分たちの反論を, 主張(A), 理由(R), 論拠(E)をもとに考える	1	2	3	4
12	自分たちの立論を, 主張(A), 理由(R), 論拠(E)のつながりを意識し考える	1	2	3	4
13	自分たちの立論と反論中の理由(R)や論拠(E)の情報に不足していないかを考える	1	2	3	4

■表6: ジャッジスキルに関する質問紙調査

1	英語の流暢性ではなく, 議論の中身をもとに判定を考える	1	2	3	4
2	両者の一部分ではなく試合全体のやり取りから, 判定を考える	1	2	3	4
3	個人的な考えや知識を判定材料にせず, 試合の議論内容から判定を考える	1	2	3	4
4	立論・反論中の主張(A), 理由(R), 根拠(E)のつながりが適切であるかを考える	1	2	3	4
5	反論が, 立論に対する適切な反論になっているかを考える	1	2	3	4
6	立論・反論中の主張(A), 理由(R), 論拠(E)を区別する	1	2	3	4
7	立論・反論中の理由(R)や論拠(E)に誤りがないかを考える	1	2	3	4
8	立論・反論中の理由(R)や論拠(E)の情報に不足がないかを考える	1	2	3	4
9	両者の立論に出されたメリット・デメリットが本当に生じるかを考える	1	2	3	4
10	肯定側から出されたメリットの重要性, 否定側から出されたデメリットの深刻性を比較し, 判定を考える	1	2	3	4

■表7: 教師によるフィードバックに関する質問紙調査

1	先生からのフィードバックは、立論・質疑応答・反論が適切につながっているかを確認する上で分かりやすい	1	2	3	4
2	先生からのフィードバックは、肯定派・否定派の主張の相違点を確認する上で分かりやすい	1	2	3	4
3	先生からのフィードバックは、立論・反論中の主張 (A)、理由 (R)、論拠 (E) が適切につながっているかを確認する上で分かりやすい	1	2	3	4
4	先生からのフィードバックは、論題や議論の中身について多角的な視点から考える上で分かりやすい	1	2	3	4
5	先生からのフィードバックは、立論・反論中の主張 (A)、理由 (R)、論拠 (E) を区別する上で分かりやすい	1	2	3	4
6	先生からのフィードバックは、立論・反論中の理由 (R) や論拠 (E) の情報が十分であるかを確認する上で分かりやすい	1	2	3	4
7	先生からのフィードバックは、メリット・デメリットが本当に生じるかを考える上で分かりやすい	1	2	3	4
8	先生からのフィードバックは、メリット・デメリットを比較する上で分かりやすい	1	2	3	4

教師からのフィードバックについてどう思いましたか。

5 結果

5.1 ディベートスキル

次頁の表8は、フィードバック別ディベートスキル事前・事後調査の平均値の増減を表したものである。0は事前・事後調査の平均値に差がないことを表し、+は増加を、-は減少を表している。「5. 相手の立論中の主張 (A)、理由 (R)、根拠 (E) のつながりが適切であるかは考えない」については、4を1に、3を2に、2を3に、1を4に処理したものを表している。グラフィックオーガナイザー・フィードバックでは全てのスキルにおいて、フローシート・フィードバックについては、「8. 自分たちの立論を、主張、理由、論拠をもとに考える」「12. 自分たちの立論を、主張、理由、論拠のつながりを意識し考える」「13. 自分たちの立論と反論中の理由や論拠の情報が不足していないかを考える」の三つのスキルを除く全てのスキルにおいて、口頭フィードバックについては、「1. 相手の立論と反論中の理由や論拠に誤りがないかを考える」「4. 自分たちの立論と相手の立論の相違点を考える」「7. 相手の立場に立ち、相手からの反論を予想する」の三つのスキルにおいて、事前より事後の平均値が高かった。

次に、フィードバック群(口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバック)とディベートスキル事前・事後調査間との差を明らかにするために二要因分散分析を行った(表9)。その結果、「6. 利害関係者の立場から論題について考えてみるといったような、多角的な視点から論題について考える ($F(2, 93) = 5.48, p = .006, \text{偏}\eta^2 = .11$)」「10. 相手の立論と反論中の理由や論拠の情報が不足していないかを考える ($F(2, 93) = 4.62, p = .01, \text{偏}\eta^2 = .09$)」「11. 自分たちの反論を、主張、理由、論拠をもとに考える ($F(2, 93) = 3.27, p = .04, \text{偏}\eta^2 = .07$)」の三つのスキルにおいて、交互作用が有意であった。交互作用があった三つのディベートスキルについて、その後、単純主効果の検定を行った。交互作用が有意でなかったスキルについては、全て主効果は認められなかった。

一つ目の「6. 多角的視点からの論題考察」において、フィードバック別に事前・事後調査の単純主効果を検定したところ、グラフィックオーガナイザー・フィードバック ($F(1, 93) = 8.39, p = .005$) が有意となり、事後調査 ($M = 3.27, SD = .57$) の方が事前調査 ($M = 2.76, SD = .79$) よりも平均値が高く ($t = -2.90, d = .66$)、効果量解釈は中であった。一方で、口頭フィードバック(事前: $M = 3.41, SD = .78$;事後: $M = 3.07, SD = .71$; $F(1, 93) = 3.30, p = .07$) とフローシート・フィードバック(事前: $M = 3.12, SD = .77$;事後: $M = 3.26, SD = .71$; $F(1, 93) = 0.71, p = .40$) は、有意でなかった。

次に、二つ目の「10. 相手の立論・反論の理由・根拠の情報不足の確認」について、フィードバック別に事前・事後調査の単純主効果を検定したところ、グラフィックオーガナイザー・フィードバックにおける事前・事後調査 ($F(1, 93) = 13.12, p = .000$) に統計的に有意な差が見られ、事後調査 ($M = 3.48, SD = .57$) の方が事前調査 ($M = 2.88, SD = .74$) よりも高く ($t = -3.62, d = .80$)、効果量解釈は大であった。その一方で、口頭フィードバッ

ク(事前: $M = 3.28$, $SD = .80$;事後: $M = 3.14$, $SD = .69$; $F(1, 93) = 0.60$, $p = .44$)とフローシート・フィードバック(事前: $M = 3.09$, $SD = .90$;事後: $M = 3.35$, $SD = .73$; $F(1, 93) = 2.58$, $p = .11$)には、有意差がなかった。

最後に三つ目の「11. 三角ロジックをもとにした自分たちの反論考察」において、フィードバック別に事前・事後調査の単純主効果を検定したところ、グラフィックオーガナイザー・フィードバックにおける事前・事後調査($F(1, 93) = 7.78$, $p = .006$)は、有意となり、事後調査($M = 3.61$, $SD = .56$)の方が事前調査($M = 3.15$, $SD = .71$)よりも平均値が高く($t = -2.79$, $d = .66$)、効果量解釈は中であつた。しかし、口頭フィードバック(事前: $M = 3.55$, $SD = .51$;事後: $M = 3.41$, $SD = .60$; $F(1, 93) = 0.63$, $p = .43$)とフローシート・フィードバック(事前: $M = 3.21$, $SD = .84$;事後: $M = 3.26$, $SD = .75$; $F(1, 93) = 0.13$, $p = .72$)には有意差が認められなかった。

■表8: フィードバック別ディベートスキル事前・事後調査の平均値の増減結果

		口頭	フローシート	グラフィック オーガナイザー
1	相手の立論と反論中の理由(R)や論拠(E)に誤りがないかを考える	+	+	+
2	相手の立論と自分たちの反論の整合性を考える	-	+	+
3	自分たちの反論中の主張(A), 理由(R), 論拠(E)のつながりが適切であるかを考える	-	+	+
4	自分たちの立論と相手の立論の相違点を考える	+	+	+
5	相手の立論中の主張(A), 理由(R), 論拠(E)のつながりが適切であるかは考えない	-	+	+
6	利害関係者の立場から論題について考えてみるといったような、多角的な視点から論題について考える	-	+	+
7	相手の立場に立ち、相手からの反論を予想する	+	+	+
8	自分たちの立論を、主張(A), 理由(R), 論拠(E)をもとに考える	-	0	+
9	相手の立論中の主張(A), 理由(R), 論拠(E)の区別を考える	-	+	+
10	相手の立論と反論中の理由(R)や論拠(E)の情報が不足していないかを考える	-	+	+
11	自分たちの反論を、主張(A), 理由(R), 論拠(E)をもとに考える	-	+	+
12	自分たちの立論を、主張(A), 理由(R), 論拠(E)のつながりを意識し考える	-	-	+
13	自分たちの立論と反論中の理由(R)や論拠(E)の情報が不足していないかを考える	-	-	+

■表9: フィードバック群とディベートスキル事前・事後調査の分散分析結果

		口頭		フローシート		グラフィック オーガナイザー		F	p	偏 η^2
		事前 テスト 平均	事後 テスト 平均	事前 テスト 平均	事後 テスト 平均	事前 テスト 平均	事後 テスト 平均			
1	相手の立論と反論中の理由(R)や論拠(E)に誤りがないかを考える	3.41	3.45	3.32	3.47	3.36	3.52	0.17	.85	.00
2	相手の立論と自分たちの反論の整合性を考える	3.48	3.38	3.26	3.38	3.39	3.45	0.55	.58	.01
3	自分たちの反論中の主張(A), 理由(R), 論拠(E)のつながりが適切であるかを考える	3.55	3.34	3.03	3.29	3.24	3.42	2.14	.12	.04
4	自分たちの立論と相手の立論の相違点を考える	3.21	3.48	3.26	3.53	3.06	3.33	0.00	1.00	.00
5	相手の立論中の主張(A), 理由(R), 論拠(E)のつながりが適切であるかは考えない	1.55	1.79	1.82	1.74	1.70	1.61	1.04	.36	.02
6	利害関係者の立場から論題について考えてみるといったような、多角的な視点から論題について考える	3.41	3.07	3.12	3.26	2.76	3.27	5.48	.006	.11
7	相手の立場に立ち、相手からの反論を予想する	3.31	3.38	3.32	3.35	3.36	3.64	0.67	.51	.01
8	自分たちの立論を、主張(A), 理由(R), 論拠(E)をもとに考える	3.59	3.41	3.21	3.21	3.27	3.58	2.03	.14	.04
9	相手の立論中の主張(A), 理由(R), 論拠(E)の区別を考える	3.10	3.07	2.91	3.06	2.73	3.30	2.69	.07	.06
10	相手の立論と反論中の理由(R)や論拠(E)の情報が不足していないかを考える	3.28	3.14	3.09	3.35	2.88	3.48	4.62	.01	.09
11	自分たちの反論を、主張(A), 理由(R), 論拠(E)をもとに考える	3.55	3.41	3.21	3.26	3.15	3.61	3.27	.04	.07
12	自分たちの立論を、主張(A), 理由(R), 論拠(E)のつながりを意識し考える	3.62	3.52	3.29	3.18	3.18	3.52	2.50	.09	.05
13	自分たちの立論と反論中の理由(R)や論拠(E)の情報が不足していないかを考える	3.55	3.24	3.29	3.24	3.12	3.42	2.48	.09	.05

5.2 ジャッジスキル

表10は、フィードバック別ジャッジスキル事前・事後調査の平均値の増減を表したものである。グラフィッ

クオーガナイザー・フィードバックでは「2. 両者の一部分ではなく試合全体のやり取りから、判定を考える」を除く全てのスキルにおいて、フローシート・フィードバックについては、「3. 個人的な考えや知識を判定材料にせず、試合の議論内容から判定を考える」「5. 反論が、立論に対する適切な反論になっているかを考える」の二つのスキルを除く全てのスキルにおいて、口頭フィードバックについては、「1. 英語の流暢性ではなく、議論の中身をもとに判定を考える」「2. 両者の一部分ではなく試合全体のやり取りから、判定を考える」「3. 個人的な考えや知識を判定材料にせず、試合の議論内容から判定を考える」の三つのスキルにおいて、事前より事後の平均値が高い結果であった。その後、フィードバック群とジャッジスキル事前・事後調査間との差を明らかにするため二要因分散分析を行った結果、交互作用と主効果において有意な差を示したスキルはなかった(表11)。

■表10: フィードバック別ジャッジスキル事前・事後調査の平均値の増減結果

		口頭	フローシート	グラフィック オーガナイザー
1	英語の流暢性ではなく、議論の中身をもとに判定を考える	+	+	+
2	両者の一部分ではなく試合全体のやり取りから、判定を考える	+	+	-
3	個人的な考えや知識を判定材料にせず、試合の議論内容から判定を考える	+	-	+
4	立論・反論中の主張(A), 理由(R), 根拠(E)のつながりが適切であるかを考える	-	+	+
5	反論が、立論に対する適切な反論になっているかを考える	-	-	+
6	立論・反論中の主張(A), 理由(R), 論拠(E)を区別する	-	+	+
7	立論・反論中の理由(R)や論拠(E)に誤りがないかを考える	-	+	+
8	立論・反論中の理由(R)や論拠(E)の情報が不足がないかを考える	-	+	+
9	両者の立論中に出されたメリット・デメリットが本当に生じるかを考える	-	+	+
10	肯定側から出されたメリットの重要性, 否定側から出されたデメリットの深刻性を比較し、判定を考える	-	+	+

■表11: フィードバック群とジャッジスキル事前・事後調査の分散分析結果

		口頭		フローシート		グラフィック オーガナイザー		F	p	偏 η^2
		事前 テスト 平均	事後 テスト 平均	事前 テスト 平均	事後 テスト 平均	事前 テスト 平均	事後 テスト 平均			
1	英語の流暢性ではなく、議論の中身をもとに判定を考える	3.45	3.72	3.59	3.76	3.55	3.64	0.26	.77	.01
2	両者の一部分ではなく試合全体のやり取りから、判定を考える	3.24	3.59	3.29	3.41	3.18	3.15	0.99	.38	.02
3	個人的な考えや知識を判定材料にせず、試合の議論内容から判定を考える	3.45	3.48	3.38	3.26	3.15	3.21	0.37	.69	.01
4	立論・反論中の主張(A)、理由(R)、根拠(E)のつながりが適切であるかを考える	3.59	3.34	3.24	3.38	3.27	3.33	1.32	.27	.03
5	反論が、立論に対する適切な反論になっているかを考える	3.79	3.55	3.62	3.59	3.58	3.64	0.79	.46	.02
6	立論・反論中の主張(A)、理由(R)、論拠(E)を区別する	3.07	2.79	3.06	3.18	2.97	3.21	1.89	.16	.04
7	立論・反論中の理由(R)や論拠(E)に誤りがないかを考える	3.45	3.31	3.24	3.53	3.39	3.58	2.19	.12	.05
8	立論・反論中の理由(R)や論拠(E)の情報が不足がないかを考える	3.24	3.21	3.26	3.53	3.15	3.42	0.80	.45	.02
9	両者の立論中に出されたメリット・デメリットが本当に生じるかを考える	3.69	3.48	3.47	3.68	3.49	3.61	1.61	.20	.03
10	肯定側から出されたメリットの重要性、否定側から出されたデメリットの深刻性を比較し、判定を考える	3.59	3.24	3.44	3.53	3.30	3.49	2.59	.08	.05

5.3 教師からのフィードバックについて

教師からのフィードバックに関するフィードバック群間の差を明らかにするため、質問紙調査結果に対して二要因分散分析を行った(表12)。その結果、全ての設問において3種類のフィードバック群間に有意な差はなかった。

次に、「教師からのフィードバックについてどう思いましたか」という質問に対する自由記述回答を

■表12: 教師からのフィードバックに関する分散分析結果

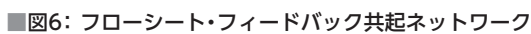
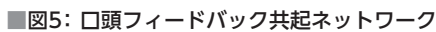
		口頭		フローシート		グラフィック オーガナイザー		p	η^2
		M	SD	M	SD	M	SD		
1	先生からのフィードバックは、立論・質疑応答・反論が適切につながっているかを確認する上で分かりやすい	3.79	0.62	3.85	0.36	3.79	0.49	.83	.00
2	先生からのフィードバックは、肯定派・否定派の主張の相違点を確認する上で分かりやすい	3.69	0.66	3.88	0.33	3.70	0.59	.26	.03
3	先生からのフィードバックは、立論・反論中の主張(A)、理由(R)、論拠(E)が適切につながっているかを確認する上で分かりやすい	3.62	0.78	3.79	0.48	3.79	0.49	.42	.02
4	先生からのフィードバックは、論題や議論の中身について多角的な視点から考える上で分かりやすい	3.59	0.73	3.62	0.49	3.70	0.53	.74	.01
5	先生からのフィードバックは、立論・反論中の主張(A)、理由(R)、論拠(E)を区別する上で分かりやすい	3.72	0.65	3.79	0.41	3.88	0.42	.47	.02
6	先生からのフィードバックは、立論・反論中の理由(R)や論拠(E)の情報が十分であるかを確認する上で分かりやすい	3.76	0.64	3.85	0.36	3.70	0.53	.46	.02
7	先生からのフィードバックは、メリット・デメリットが本当に生じるかを考える上で分かりやすい	3.59	0.83	3.68	0.48	3.70	0.47	.75	.01
8	先生からのフィードバックは、メリット・デメリットを比較する上で分かりやすい	3.55	0.83	3.62	0.65	3.70	0.53	.70	.01

KH Coder 3.Beta.03i (樋口, 2014) により分析し、テキストマイニング処理を行った。その結果、口頭フィードバックは44文, 38段落, 総抽出語数929語, 異なり語数221語, フローシート・フィードバックは、45文, 41段落, 総抽出語数1010語, 異なり語数237語, グラフィックオーガナイザー・フィードバックは、36文, 34段落, 総抽出語数618語, 異なり語数173語であった。各フィードバックの抽出語の最小出現回数を2回とし、抽出語間の関連性を示す Jaccard 係数を使用し、各語の頻出回数の違いを円の大きさの違いで示した。

図5は、口頭フィードバック共起ネットワークを示している。最も大きな円である「自分」とつながる大きな円として、「反論」と「考える」がある。「自分」「反論」に関連する回答は、「自分たちの立論や反論がどのように捉えられたのか」「自分たちの立論が反論によってどうなったのか」について理解できるという内容であった。「自分」とのもう一つのネットワークである「自分」「考える」は、「自分が考えきれないような視点からのフィードバックで、これからどのようにジャッジを行えば良いのか参考になる」という内容であった。円の大きさでその他目を引くネットワークとして、「一つ一つ」「主張」「説明」「丁寧」から成るネットワークと、「論理」「重要」「ジャッジ」「仕方」から成るネットワークがある。前者は「一つ一つの主張や反論、質問を丁寧に振り返り、説明してくれる所が分かりやすかった」、後者は「論理性と重要性という観点によるジャッジの仕方が理解できた」という内容であった。

次に、図6はフローシート・フィードバック共起ネットワークを示しており、最も大きな円である「立論」は「質問」「理由」「根拠」と結びついている。その中身は、「立論(主張)と理由、根拠はつながっているのか」「質問は的確にできているのか」を確認する上で分かりやすかったという内容であった。また、フローシート・フィードバック特有のネットワークに「相手」「全体」「的確」「確認」「攻撃」から成るネットワークがあった。「自分たちの反論が的確に相手の立論を反論することができているのか、全体の流れの中で、フローシートの列間の内容を比較することにより確認できて良かった」という回答であった。

最後に、図7はグラフィックオーガナイザー・フィードバック共起ネットワークを示している。最も大きな円である「反論」は「後で」「前」「比較」「難しい」と結びついている。その内容は、「反論が主張(立論)のどこに反論しているのか、反論後の主張(立論)の弱まり方を反論の前と後でとても比較しやすかった」というものであった。グラフィックオーガナイザー・フィードバックに特徴的なネットワークに、「矢印」「色」「変わる」「良い」「簡潔」から構成されるネットワークがあった。その内容は、「簡潔に文がまとめられていて、話の流れの把握が矢印や色でわかりやすかった」「論理性や重要性への評価の変わり方が分かりやすく良い」といった回答であった。また、「聞く」「ディベート」から成るネットワークでは、聞いているだけでは理解が難しいディベートの内容が、図示によりその理解が容易になった旨が述べられていた。



なる。先行研究では、立論や反論中の主張を支持する理由・根拠に含まれる十分な情報の存在が、その質を決める上での重要な要因の一つであるとされている(e.g., Stapleton & Wu, 2015)。本研究におけるグラフィックオーガナイザーの各枠内には、同じ言葉の重複やつなぎ言葉などが排除された鍵となる最低限の情報のみが含まれており、このことは、参加者に情報の焦点化と整理を容易にし、アーギュメントの質を決定する情報不足の確認に対する意識を鋭敏にしたのではないかと考えられる。

最後に、「11. 三角ロジックをもとにした自分たちの反論考察」に関して考察したい。ライティングの多くの先行研究では、反論とそれに対する反駁はアーギュメントの説得性を増すとされている(e.g., Wolfe, Britt & Butler, 2009)が、日本人高校生が反論と反駁をライティングに導入することの難しさも指摘されている(坂口, 2021)。ディベートという高次思考が求められるスピーキング活動においても、日本人高校生にとって反論考察は容易なものではないと考えられる。反論考察には、多角的視点による反論箇所の発見、そして反論内容の考察という手順が必要である。ディベートでは、理由、論拠、主張と理由・論拠のつながりの3箇所いずれかに反論を加える。参加者は、三角ロジックという論理の仕組みを端的に表したグラフィックオーガナイザーを参考にすることにより、理由、論拠、主張の3要素の区別を行ったり、各要素の論理的つながりを確認することが可能となり、反論箇所の発見と三角ロジックを基盤とした反論内容の考察が容易になったと思われる。さらに、松下・前田・田中(2022)は、反論は問題に対する異なる視点や視野の獲得であるとしており、「6. 多角的視点からの論題考察」の向上に伴い、反論考察においても多角的視点の獲得が、反論考察を促した可能性があると考えられる。

その一方で、統計上の結果からは、フィードバックによるディベートスキルへの効果が見られなかったスキルもあった。以下、これらのスキルを6つに分類したい。(1)「1. 相手の立論と反論中の理由や論拠に誤りがないうかを考える」といった「情報の真偽の判断」、(2)「2. 相手の立論と自分たちの反論の整合性を考える」「4. 自分たちの立論と相手の立論の相違点を考える」といった「異なる主張間の整合性と相違点の確認」、(3)「3. 自分たちの反論中の主張、理由、論拠のつながりが適切であるかを考える」、(5. 相手の立論中の主張、理由、論拠のつながりが適切であるかは考えない(反転項目)」、(12. 自分たちの立論を、主張、理由、論拠のつながりを意識し考える」といった「三角ロジックの各要素のつながりへの意識」、(4)「7. 相手の立場に立ち、相手からの反論を予想する」といった「相手からの反論予想」、(5)「9. 相手の立論中の主張、理由、論拠の区別を考える」といった「相手の立論中の三角ロジックの各要素の区別」、(6)「8. 自分たちの立論を、主張、理由、論拠をもとに考える」「13. 自分たちの立論と反論中の理由や論拠の情報が不足していないかを考える」といった「自分たちの立論・反論の熟考」である。これらのスキル向上には、自らの思考の癖への気づき、主張の中身の吟味、分析、比較、検証、評価といった高次な思考が求められる。(1)「情報の真偽の判断」には、相手の主張を鵜呑みにしない情報の客観的分析、評価、(2)「異なる主張間の整合性と相違点の確認」には、複数情報の整理と比較、検証、(3)「三角ロジックの各要素間のつながりへの意識」と(5)「相手の立論中の三角ロジックの各要素の区別」には、暗黙の前提となり省略されやすい論拠の存在への気づきや、中身が不十分な論拠を自らの推論で内容を補い、解釈してしまう傾向にある自らの思考の癖への気づき、(4)「相手からの反論予想」には、議論の流れの網羅的な振り返りと、自分の立論の妥当性の検証、評価、(6)「自分たちの立論・反論の熟考」には、客観的思考が欠如し、自己中心的な思考をしてしまうという思考の癖への気づきが必要であると考えられる。統計上、本研究で実施した口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバックの方法や頻度では、こうした思考の癖への気づき、主張の吟味、分析、比較、検証、評価に必要な高次思考を促すには十分な効果がなかったと推察される。しかし今後は、参加者へのインタビューなどを通し、数値的結果からは見えない3種類のフィードバックが高次思考へ与える効果を探る必要があるだろう。

6.2 ジャッジスキル

RQ2の異なるフィードバックによるジャッジスキルへの効果においては、フィードバック群とジャッジスキル事前・事後調査間の交互作用において、有意な差を示したスキルはなかった。ディベートスキルの中にはフィードバックによる効果が見られたスキルもあったため、フィードバック効果にはディベートスキルとジャッジスキル間では差がある結果となった。これは、ディベーターに比べジャッジへの認知負荷が高く、本研究で行った3種類のフィードバックは、ディベーターとジャッジの認知負荷の差を埋めるには十分でなかったことが影響していると思われる。

この両者の認知負荷の差異を生む背景には、事前把握できる情報量と、果たすべき役割の違いがあると考えられる。チームプレイであるディベートはその特徴から、ディベーターは、自分サイドの立論や反論の内容を事前に把握することは可能であるが、ジャッジは、賛成・反対両サイドの議論内容を試合で初めて聞くため、事前把握できる情報はほとんどない。そのため、ジャッジは、試合において議論の流れやその中身を理解する必要があり、ディベーターに比べその認知負荷は高い。また、ジャッジは、ディベーターと異なり、提示された両サイドの主張を比較、検証し、最終的に勝敗を決定するという役割を担っている。判定決断に至るまで、ジャッジは、賛成・反対両サイドの各議論内容をフォローし、両者の相違点や対立点を考え、論理性や重要性を熟考しなければならず、その認知負荷は高い。このことから、事前把握可能な情報量と役割において、ディベーターとジャッジ間には認知負荷の差が生じると考えられるが、本研究で実施した2回の教師ジャッジによるフィードバックでは、これらの認知負荷の差を埋めるには十分でなかったと思われる。

6.3 教師からのフィードバック

RQ3の各フィードバックによる教師ジャッジの思考過程の気づきへの効果については、全ての設問において3種類のフィードバック間に有意な差は認められなかった。しかし、KH Coderによる自由記述回答の分析結果からは、各フィードバックが高校生英語ディベート初心者の教師ジャッジの思考への気づきに寄与したことが伺えた。具体的には、参加者は口頭フィードバックを通し、自分たちの立論や反論についての客観的評価、ジャッジの仕方、自分とは異なるジャッジの視点、ジャッジの評価観点について丁寧な説明を受けたという印象を持っていることが分かった。フローシート・フィードバックについては、フローシートの支援により、三角ロジックの各要素のつながりや質問の適切さの確認が容易になり、また、フローシートの列間の内容比較により、立論と反論の整合性の確認を試合全体の流れの中で行うことができたことが明らかとなった。グラフィックオーガナイザー・フィードバックについては、矢印、色分けした枠を用いた図示は、参加者に、聴覚情報として得た議論内容の視覚情報による確認、議論の流れの理解、反論箇所と反論後の相手立論の弱体化への効果についての理解を促すことが分かった。

7 結論と今後の課題

本研究は、高校生英語ディベート初心者を対象に、口頭フィードバック、フローシート・フィードバック、グラフィックオーガナイザー・フィードバックの3種類のフィードバックを行い、ディベートスキル、ジャッジスキル、教師ジャッジの思考過程の気づきへの効果を検証した。その結果、グラフィックオーガナイザー・フィードバックにより、「多角的視点からの論題考察」「相手の立論・反論の理由・根拠の情報不足の確認」「三角ロジックをもとにした自分たちの反論考察」の三つのディベートスキルの伸長が促された。ジャッジスキルについては、フィードバック群と事前・事後調査間に交互作用が見られなかったことから、統計的結果からは、本研究で実施した各フィードバックの方法や頻度は、高校生英語ディベート初心者のジャッジ

スキルを向上させるには十分でなかった。高校生英語ディベート初心者の教師ジャッジの思考への気づきについては、フィードバック間に有意な差はなかった。しかし、KH Coderによる自由記述回答の分析結果からは、各フィードバックは、高校生英語ディベート初心者がジャッジをする上で必要なロジックや主張の中身の吟味、異なる主張間の整合性の確認、議論展開の理解を助け、教師ジャッジの視点やジャッジ方法についての気づきを促す可能性が示唆された。

しかし、本研究には、いくつかの課題も残る。一つ目に、本研究でフィードバックによる向上が見られなかったディベートスキルやジャッジスキルの伸長を図る上で、フィードバックは有効なのかどうかを明らかにする必要がある。そのため、ピア・フィードバックなどの異なるフィードバックを用いる、あるいは本研究で用いたフィードバックの回数や頻度を変えるなどし、今後、検証する必要がある。次に、ディベートスキルにおいて、「10. 相手の立論・反論の理由・根拠の情報不足の確認」には交互作用が見られたが、「13. 自分たちの立論と反論の理由や論拠の情報不足の確認」には、主効果も交互作用も見られなかった。また、「11. 三角ロジックをもとにした自分たちの反論考察」には交互作用が見られたが、「8. 三角ロジックをもとにした自分たちの立論考察」においては、主効果も交互作用も見られなかった。このことから、なぜフィードバックは、自分たちの主張よりも相手の主張の情報不足の確認に効果があるのか、また、三角ロジックをもとにした考察においては、立論よりも反論考察に効果があるのか、今後、更なる検証が必要である。今後の研究課題としたい。

謝辞

本研究の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、選考委員の諸先生方、参加者の皆様に感謝申し上げます。中でも、横浜国立大学 斉田智里先生には、貴重なご助言と有益なご指導をいただきました。この場を借りて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

- Akerman, R., & Neale, I. (2011). *Debating the evidence: An international review of current situation and perceptions*. CfBT Education Trust [in association with] English-Speaking Union.
- Benesse. (2016). 『中高の英語指導に関する実態調査2015』. ベネッセ教育総合研究所. https://berd.benesse.jp/up_images/research/03_Eigo_Shido.pdf より (2023年12月29日閲覧)
- Ellis, E., & Howard, P. (2007). Graphic organizers: Power tools for teaching students with learning disabilities. *Current Practice Alerts*, 13(1), 1-4.
- 藤岡信勝・上條晴夫. (1996). 『中高生のためのやさしいディベート入門』. 学事出版.
- 福澤一吉. (2018). 『新版 議論のレッスン』. NHK出版.
- Hanson, J. (1997). Justifying decisions: Good ballots give good reasons. *Southern Journal of Forensics*, 2, 87-105.
- 樋口耕市. (2014). 『社会調査のための計量テキスト分析—内容分析の継承と発展を目指して—』. ナカニシヤ出版.
- 市川裕理. (2018). 「「即興で話す」力につなげるディベート実践」. 『中部地区英語教育学会紀要』, 47, 149-156.
- 櫻尾文雄. (2011). 「高校生のディベート指導を通しての成果と課題」. 『四国英語教育学会紀要』, 31, 69-78.
- 加藤彰. (2020). 『即興型ディベートの教科書: 東大で培った「瞬時に考えて伝えるテクニック」』. あさ出版.
- 河野周. (2021). 『中学・高校英語ディベート入門』. 三省堂.
- 小林良裕. (2019). 『Book 3: FOR CLASSROOM USES 英語ディベーター高校授業用テキスト(教員用) —』. NextPublishing Authors Press.
- 小林良裕. (2021). 「即興型英語ディベートでのジャッジ間の評価の違い: 高校生全国大会決勝の事例分析から」. 『ディベートと議論教育: ディベート教育国際研究会論集』, 3, 35-50.
- 是澤克哉. (2002). 「良いジャッジとはどうあるべきか」『ディベートフォーラム』. http://www.debateforum.org/2002/2d08_koresawa.pdf より (2023年12月29日閲覧)
- 久保健治. (2017). 「日本語ディベートにおけるジャッジ教育の方法論: ジャッジインターン副審養成講座の実践報告」. 『ディベートと議論教育: ディベート教育国際研究会論集』, 1, 62-74.
- 松本茂. (1996). 『頭を鍛えるディベート入門: 発想と表現の技法』. 講談社.
- 松本茂・鈴木健・青沼智. (2008). 『英語ディベート 理論と実践』. 玉川大学出版部.
- 松下佳代. (2021). 『対話型論証による学びのデザイン: 学校で身につけてほしいたった一つのこと』. 勁草書房.

引用文献

- 松下佳代・前田秀樹・田中孝平. (2022). 『対話型論証ですすめる探究ワーク』. 勁草書房.
- 文部科学省. (2018). 『高等学校学習指導要領解説外国語編・英語編』. 開隆堂出版.
- 茂木秀昭. (2001). 『ザ・ディベート: 自己責任時代の思考・表現技術』. 筑摩書房.
- 茂木秀昭. (2013). 『「60分」図解トレーニング ロジカル・シンキング』. PHP 研究所.
- 内藤真理子・西村由美. (2018). 『大学生のためのディベート入門: 論理的思考を鍛えよう』. ナカニシヤ出版.
- 中川智皓. (2019). 「パラメンタリーディベートー論点の可視化」. 『システム/制御/情報』, 63(4), 170-175.
- 中井弘一. (2022). 「中学校英語教科書に見るディベート力の育成ー論理的な思考力・判断力・表現力を育む英語授業に向けてー」. 『大阪女学院大学紀要』, 18, 157-170.
- 中野美香. (2010). 『大学1年生からのコミュニケーション入門』. ナカニシヤ出版.
- Praveen, S. D., & Rajan, P. (2013). Using Graphic Organizers to Improve Reading Comprehension Skills for the Middle School ESL Students. *English language teaching*, 6(2), 155-170.
- 坂口寛子. (2021). 「高校生を対象とした反論と反駁を含む英語論証文指導」. 『EIKEN BULLETIN』, 33, 132-153.
- 坂口寛子. (2022). 「『論理・表現』における英語ディベート指導に関する考察ーディベートに関する調査結果に焦点をあてて」. 『LET Kyushu-Okinawa BULLETIN』, 22, 27-36.
- 佐藤紀美子. (2011). 「ディベートのフィードバックにおける日本語学習者の気づきの分析」. 『日本語教育研究/長沼言語文化研究所編』, 57, 94-108.
- 佐藤紀美子. (2015). 「ディベートの発話に対する日本語母語話者と日本語学習者の評価に関する一考察」. 『同志社大学日本語・日本文化研究』, 13, 97-114.
- 清水裕士. (2016). 「フリーの統計分析ソフトHAD: 機能の紹介と統計学習・教育, 研究実践における利用方法の提案」『メディア・情報・コミュニケーション研究』, 1, 59-73.
- Stapleton, P., & Wu, Y. A. (2015). Assessing the quality of arguments in students' persuasive writing: A case study analyzing the relationship between surface structure and substance. *Journal of English for Academic Purposes*, 17, 12-23.
- 鶴田清司. (2017). 『論理的思考力・表現力を育てる三角ロジック』. 図書文化社.
- Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication*, 26(2), 183-209.
- 山崎康司. (2011). 『入門考える技術・書く技術: 日本人のロジカルシンキング実践法』. ダイアモンド社.
- 矢野善郎. (2020, October 25). *Judging Debate* (also) *Makes Friends!*. HenDA Judging Seminar. [PowerPoint slides]. http://henda.global/bc_rma/wp-content/uploads/2020/12/HEnDA-Judge-Seminar-20201025.pdf より (2023年12月29日閲覧)
- 保田幸子. (2024). 『「書く力」の発達ー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ー』. くろしお出版.
- 安井省侍郎. (1994). 『初心者のためのディベート Q&A』. Independently published.
- 横山雅彦. (2006). 『高校生のための論理思考トレーニング』. 筑摩書房.

第36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

明示的知識を外化して正確性の向上を 目指す指導の効果

— 長期的なランゲージング・エピソードの分析を中心に —

研究者:高木 哲也 東京都／筑波大学附属高等学校 教諭

《研究助言者:西垣 知佳子》

概要

本研究は、長期的なランゲージング活動の学習効果を検証した。日本語を母語とする高校1年生61名を対象に、約5ヶ月間にわたって30語程度の自由英作文課題に対して間接書記訂正フィードバックを与えて、その後口頭または筆記ランゲージング活動を行う処遇を8回実施した。分析結果より、英文法テストでは口頭ランゲージング群が有意に向上し、筆記ランゲージング群より高い学習効果が示唆されたが、自由英作文テストの正確性は両群共に向上しなかった。記録されたランゲージング・エピソードの字数に2群間の差はないものの、気づきレベルの高いランゲージング・エピソードの頻度数の観点では、筆記ランゲージング群の優位性が示された。さらに、学習活動として行う「思考の言語化」を肯定的に捉える学習者は、ランゲージング活動への関与が強くなる傾向が見られた。本研究より、正確性の向上を目指した効果的なライティング指導法として、口頭または筆記ランゲージング活動を選択して活用することが提案される。

1 はじめに

1.1 英語ライティング指導の意義

現行の学習指導要領は、「聞くこと」、「読むこと」、「話すこと[やり取り]」、「話すこと[発表]」、「書くこと」の4技能5領域の力を統合的にバランスよく育成すること、特に「話すこと[やり取り]」、「話すこと[発表]」、「書くこと」による発信能力を育成する必要性を示している。しかし、文部科学省による「平成29年度英語力調査結果(高校3年生)の概要」によると、高校3年生の英語力は、「話すこと」と「書くこと」は「聞くこと」や「読むこと」と比較して全体的に低く、無得点者の割合も依然として高い状況であり、4技能がバランスよく育成されていない課題がある。さらに、ベネッセ教育総合研究所による「中高生の英語学習に関する実態調査2014」は、中高生の半数以上が「英語の文を書くのが難しい」ことや「文法が難しい」ことを学習上の躓きとして捉えており、そのように回答した高校生は中学生より10%以上多いことを報告している。つまり、国内の英語教育における発信能力の育成、特に「書くこと」に関する教師の支援の充実がより一層求められている。

1.2 英語ライティング指導における「振り返り活動」の重要性

英語ライティング指導において、正確性の向上は重要な目標の一つである。特に、外国語学習者が自身

の誤りに気づき、それを素早く修正できる高いメタ認知力を身につけることは、「自立した書き手を育てること」と密接に関わる(広瀬, 2014, p.132)。つまり、学習者自身が自分の書いた英文を見直し、誤り等を訂正する「振り返り活動」を継続することは、ライティング力の向上に効果的な学習法の一つである(駒井, 2024; 松下, 2024)。さらに、「自律した学習者」の育成を目指して「自らの学習状況を把握し、学習の進め方について試行錯誤するなど自らの学習を調整しながら、学ぼうとしているか」(国立教育政策研究所, 2020, p.10)を評価する手立てとしても「振り返り活動」の重要性が再認識されている(吉川, 2024)。

1.3 ランゲージングとは

第二言語習得研究において、学習者が言語面に関して話したり書いたりして振り返る行為は「ランゲージング(languaging)」と呼ばれている(Swain, 2006)。ランゲージングは「我々が第二言語を学び、上級レベルまで到達するための一つの手法であり(“one of the ways we learn a second language to an advanced level.”)(Swain, 2006, p.96)、学習者の深い思考を促し、様々な気づきを与える機会となるため、多くの実証研究によって、特に言語面の正確性の向上に寄与すること」が報告されている(e.g., Li et al., 2023; Niu & Li, 2017; Suzuki & Storch, 2020; Swain & Watanabe, 2013; 高木, 2023b; 布川, 2018)。ランゲージングの形態は、口頭ランゲージング(oral languaging: 以降, OL)と筆記ランゲージング(written languaging: 以降, WL)があるが、国内の教室現場でOLやWLを実践した研究はその数に限りがある(e.g., Ushiyama, 2022; 市川, 2022; 今澤, 2024; 江下, 2020; 佐藤, 2020; 高木, 2022a, b, 2023a, 2024)。

1.4 形態の異なるランゲージングの学習効果を比較検証した先行研究

Wang (2023) は、中国語を母語とする大学生75名を対象に、ランゲージングの形態(OLとWL)と指示文の有無(例: どのような形式を書いたか。教師からどのような提案があったか。)を条件に統制群を含めて5つの群を設定した。統制群を除く4つの群は、200語程度の自由英作文課題に対して、明示的に誤りを訂正した直接書記訂正フィードバック(direct written corrective feedback: 以降, DWCF)が与えられた後にランゲージングを行った。直後ポストテスト・遅延ポストテスト(1ヶ月後)として行った誤り修正テストの結果を用いて学習効果を分析した。結果、遅延ポストテストにおける誤り修正の成功率はWL群がOL群より高く、WLの長期的な学習効果が示された。さらに、指示文が与えられたWL群は、指示文が与えられなかったWL群の約2倍の誤り修正成功率を示したため、指示文の有無がランゲージングの学習効果に大きな影響を与えることが示唆された。

Peng et al. (2023) は、中国語を母語とする中学生81名(CEFR A1-A2レベル)を対象に、個人で取り組むWL群とペアで取り組むOL群を設定した。70-100語程度の自由英作文課題に言語面・内容面・構成面に関するDWCFが与えられた後にランゲージングを行う活動を3回実施した。ポストテストとして自由英作文課題の書き直しを行い、参加者が口頭や筆記で言語面に関する考えを記録したランゲージング・エピソード(languaging episode; 以降, LE)の内容を含めて学習効果を検証した。結果、個人で取り組むWL群はDWCFに対する修正率がより高く、ペアで取り組むOL群はDWCFが与えられた箇所だけでなく、より広範囲の観点で振り返る傾向が示された。LEの内容から、ペアで取り組むOL群は個人で取り組むWL群よりも、深い認知処理を行う傾向が見られたが、書き直しにおける正確性に2群間の差は見られなかった。ランゲージングの形態の違いは、学習者の書記訂正フィードバック(written corrective feedback: 以降, WCF)に対する異なる行動や認知プロセスを導く可能性が示唆された。

Ushiyama (2022) は、日本語を母語とする中学1年生53名を対象に、誤りの箇所のみ指摘する(1)間接書記訂正フィードバック(indirect written corrective feedback: 以降, IWCF) + languaging群と(2)DWCF群を設定した。ターゲットとなる英文法を三人称単数現在形に設定して、プレテストとして「私の家族」について、自由英作文およびスピーキングテストを実施した。WCFが与えられた後、(1)はペアで

対話をしたり(OL)気がついたことを書いたりして(WL)誤り箇所を修正するよう指示され、(2)は各自で声に出さずに修正箇所を確認するよう指示された。(1)に関して、OLやWLそれぞれに取り組む程度は生徒に委ねられており、結果としてOLとWLどちらかを選択した参加者と両方に取り組んだ参加者がいた。10日後にプレテストと同じトピックで自由英作文およびスピーキングテストを実施し(ポストテスト)、1ヶ月後と8ヶ月後に実施した確認テストの内、三人称単数現在形の得点を抽出して長期的な学習効果を検証した。結果、ポストテストにおける正確性は、自由英作文およびスピーキングの両技能において両群共に有意に向上した。さらに、(1)IWCF + languaging群の生徒の内、OL + WLに取り組んだ学習者が最も高い学習効果を示しただけでなく、言語化すること自体が難しい学習者が友人の説明を聞くことで学習効果につながった可能性も示唆された。

1.5 研究課題

先行研究では、OLとWLの学習効果の違いや形態の異なるランゲージングで記録されるLEの質の違いが検証されているものの、国内の教室現場において個人で取り組むOLとWLを長期的に実践した研究は見当たらない。英語ライティング指導における言語面の正確性の向上を目指した指導法として、さらに「自律した学習者」の育成に貢献し得る指導法としてのランゲージングを長期的に実践する場合、口頭と筆記の形態の違いによって、どのように異なる学習効果や認知処理を伴うか検証する意義は大きいと考える。よって本研究では、日本語を母語とする高校生を対象に、以下の3つの研究課題を設定した。

研究課題 1 長期的にOLまたはWLに取り組むことで言語面の正確性は向上するか

研究課題 2 LEは口頭と筆記で量的・質的に違いがあるか

研究課題 3 【研究課題 2】で違いが認められた場合、英語ライティングに対する考えとLEの頻度や質はどのような関係があるか

2 研究方法

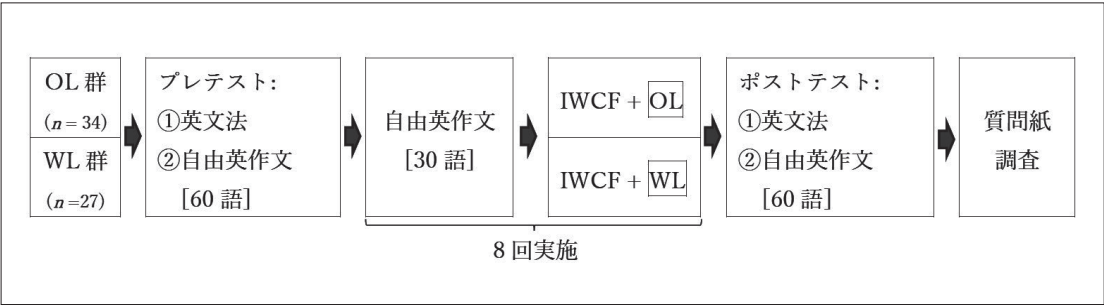
2.1 参加者

日本語を母語とする高校1年生75名を対象に、初めに本研究の目的と研究参加者の権利やプライバシーの保護、研究内容の公表について口頭で説明した。研究協力の同意を書面で得た生徒の内、全てのテスト・質問紙調査に参加し、8回実施した自由英作文課題の内、6回以上提出した生徒61名を分析対象とした。

2.2 実践手順

図1は本研究の実践手順を示す。令和5年10月から令和6年2月の英語コミュニケーションⅠの授業内で、教科書の内容に関する30語程度の自由英作文課題を8回実施した。その後、IWCFを付記して課題を返却した際に、「添削された箇所はなぜ誤りなのか、考えて説明する」というランゲージング活動を、OL群(34名)とWL群(27名)ともに10分間で実施した。両群共にGoogleフォームを利用してLEを提出した。つまり、OL群(34名)は口頭で録音したデータを添付して提出し、WL群(27名)はタイピング書字で回答して提出した。プレテスト・ポストテストとして、英文法テスト(西垣ほか、2024を基に筆者が改訂)と60語程度の自由英作文テストを実施した。英文法テストは中学校英文法事項に関する空所補充の記述式問題(43題)

であり、参加者は Google フォーム を利用して 20 分間で回答した (資料 1)。自由英作文テストは 15 分間で
 行い、プレテストは「尊敬する人」、ポストテストは「信頼する人」という類似のトピックに対して取り組んだ。
 ポストテスト後に英語ライティングに関する質問紙調査 (6 件法) を行った (資料 2)。なお、参加者はランゲ
 ージング活動を複数回経験しているが、ランゲージングの経験値が学習効果に影響を与える可能性が示さ
 れている点を鑑み (e.g., Ishikawa, 2015; Kazemi et al., 2022), 1-2 回目のランゲージング活動後に記録
 された LE を複数取り上げ、クラス全体で共有する指導を行った。



■ 図1: 実践手順

2.3 分析方法

プレテスト・ポストテストで実施した英文法テストは1問1点で採点した (43 点満点)。なお、大文字・小文
 字の表記ミスや、不要なスペースを入れた回答は不問とした。プレテスト・ポストテストで実施した自由
 英作文テストの評価は、英語指導歴 20 年以上のアメリカ出身ネイティブ教員 1 名に対して、筆者が DWCF
 例を示したうえで、言語面の正確性に焦点を当てた同様の DWCF を付記するよう依頼した。生徒それぞれの
 自由英作文の誤り総数を総語数で割り、平均語数を掛けて算出した「正規化した誤り数値」を用いて分
 析した (Suzuki, 2012; 図 2)。英語ライティングに関する質問紙調査の回答結果は、探索的因子分析を行った。

正規化した誤り数値

=

自由英作文の誤り総数

自由英作文の総語数

×

自由英作文の平均語数

■ 図2: 正規化した誤り数値の算出方法 (Suzuki, 2012)

【研究課題 1】「長期的に OL または WL に取り組むことで言語面の正確性は向上するか」に対しては、プ
 レテスト・ポストテスト (英文法・自由英作文) それぞれの得点に 2 群間の有意差があるか検証するために、
 指導法 (OL 群・WL 群) × テスト時期 (プレ・ポスト) の二元配置分散分析を行った (JASP 0.18.3.0 を使用)。
 英文法テストでは天井効果が見られたため、追加分析として、プレテストにおいて正答率が 8 割未満だっ
 た参加者 16 名に焦点を当てて、プレテスト・ポストテスト (英文法) の得点を用いて 2 群間の差がないカ
 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実施した (JASP 0.18.3.0 を使用)。さら
 に、英文法テストにおいて 8 割未満の正答率であった関係詞 [who(m)/which/that] に注目し、自由英作
 文テストにおける関係詞 who(m) と which, that の使用頻度、その中で正しい使用回数と誤った使用回数
 を AntConc (ver. 4.2.4) を使用して数えた。

【研究課題 2】「LE は口頭と筆記で量的・質的に違いがあるか」に対しては、OL 群と WL 群それぞれの LE

総字数に2群間の有意差があるか検証するために、ノン・パラメトリック検定のマン・ホイットニー検定を実施した(JASP 0.18.3.0を使用)。さらに、気づきレベルと記述内容の正誤を条件に5つに分類したLE数を数えた後(表1)、それぞれの分類のLE数に2群間の有意差が見られるか χ^2 検定を行った(js-STAR XR+release 2.1.2 jを使用)。

■表1: 自由英作文課題の解答例,およびLEの5分類とその例(原文ママ)

LE 分類	自由英作文課題の解答例とLE 例
理由を伴った 正しい	<i>People enjoy <u>to talk</u> to local friends in winter festivals.</i> → enjoy の後に to talk としてしまいました。enjoy の後は動名詞じゃないといけないので, talking
理由を伴った 誤った	<i><u>Because</u> there is so dangerous....</i> → because は文頭では用いられないので, since などを用いて説明する。
単純な気づき 正しい	<i>If “chickenosaurus” was created, I <u>want</u> to take care of it.</i> → would like にするところを want にしてしまった。
単純な気づき 誤った	<i>If I <u>chance</u> to go to Cambodia,...</i> → chance を happen にすると動詞としていい。
半信半疑	<i>..., so I can’t <u>challenge</u> to go to Cambodia.</i> → 分かりませんでした。

注. LEの5分類は, 鈴木(2008)や高木(2024)の分類基準を基に筆者作成。下線は筆者が与えたIWCFを示す。

【研究課題 3】「【研究課題 2】で違いが認められた場合、英語ライティングに対する考えとLEの頻度や質はどのような関係があるか」に対しては、【研究課題 2】で得られたLEの頻度や質の違いと質問紙調査を探索的因子分析した結果得られた因子との関連を検証するために、スピアマンの順位相関分析を行った(JASP 0.18.3.0を使用)。

3 結果と考察

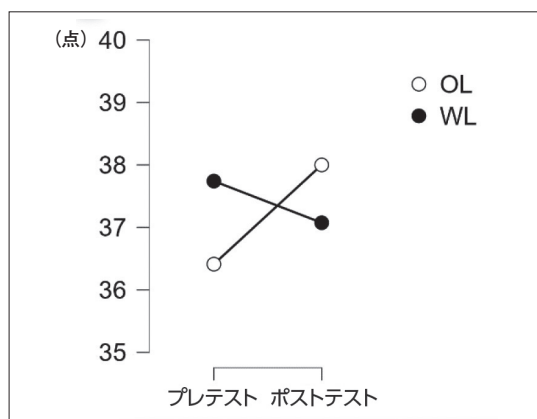
3.1 研究課題 1:長期的にOLまたはWLに取り組むことで言語面の正確性は向上するか

表2はプレテスト・ポストテスト(英文法・自由英作文)の得点の記述統計を示す。プレテスト・ポストテスト(英文法)のクロンバックの信頼性係数は $\alpha = .89$ と $\alpha = .88$ であり、十分望ましい信頼性を示した。プレテスト(英文法・自由英作文)の得点を用いて2群間の差がないか対応のない t 検定を実施したところ、有意差は見られず、効果量はそれぞれ小程度(Cohen, 1988)であった[それぞれ $t(59) = -0.88, p = .40, d = .22$; $t(59) = 0.25, p = .80, d = .06$]。よって、プレテストを実施した時点で、2群の英文法力と英作文の産出力に差はなかった。

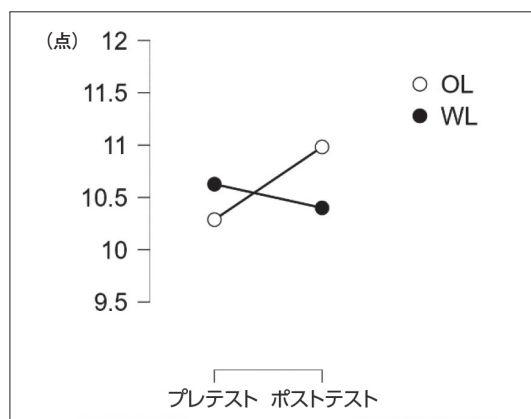
■表2: プレテスト・ポストテスト(英文法・自由英作文)の得点の記述統計

指導法	英文法				自由英作文			
	プレテスト		ポストテスト		プレテスト		ポストテスト	
	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差
OL群($n = 34$)	36.41	6.89	38.00	5.80	10.29	6.18	10.98	3.67
WL群($n = 27$)	37.74	4.67	37.07	5.02	10.63	3.77	10.40	2.64

注. 英文法テストは43点満点, 自由英作文の得点は「正規化した誤り数値(図2)」である。



■図3: 各水準の平均値のプロット (英文法)



■図4: 各水準の平均値のプロット (自由英作文)

プレテスト・ポストテストの得点に2群間の有意差があるか検証するために、指導法(OL群・WL群)×テスト時期(プレ・ポスト)の二元配置分散分析を行った結果(表3)、英文法テストでは交互作用が有意であったため $[F(1, 59) = 8.00, p < .01, \eta_p^2 = .01]$ 、単純主効果の検定を行った。その結果、OL群はプレーポスト間で得点が有意に上がったことが中程度の効果量(Cohen, 1988)と共に示されたが $[t(33) = -2.75, p = .01, d = .47]$ 、WL群はプレーポスト間で得点が下がったものの、その得点間には有意差が認められず、効果量も小程度(Cohen, 1988)だった $[t(26) = 1.28, p = .21, d = .25]$ 。

■表3: 二元配置分散分析の結果 (英文法テスト)

ケース	タイプⅢ平方和	自由度	平均平方	F	p	η_p^2
テスト時期	6.39	1	6.39	1.34	.25	.02
テスト時期×指導法	38.26	1	38.26	8.00	.01	.01
誤差	282.12	59	4.78			
指導法	1.22	1	1.22	0.02	.89	< .001
誤差	3617.16	59	61.31			

一方、表4が示すとおり、自由英作文テストでは交互作用が有意でなく $[F(1, 59) = 0.46, p = .50, \eta_p^2 = .01]$ 、テスト時期と指導法それぞれの主効果も見られなかった [それぞれ $F(1, 59) = 0.12, p = .73, \eta_p^2 = .002$; $F(1, 59) = 0.02, p = .89, \eta_p^2 < .001$]。つまり、8回のランゲージング活動を通して、英文法テストで測定される言語面の正確性はOL群の優位性が示唆されたが、より自由な産出形式である自由英作文テストにおいては両群共に向上しなかった。

■表4: 二元配置分散分析の結果 (自由英作文テスト)

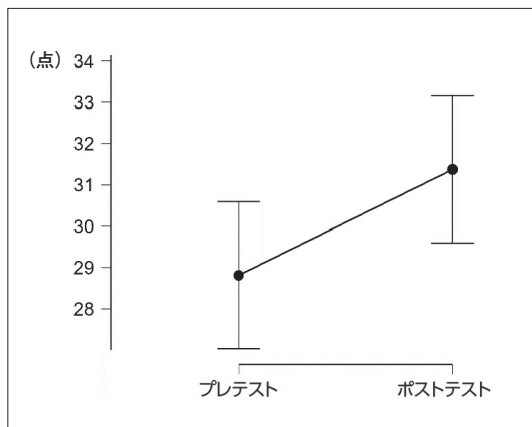
ケース	タイプⅢ平方和	自由度	平均平方	F	p	η_p^2
テスト時期	1.65	1	1.65	0.12	.73	.002
テスト時期×指導法	6.39	1	6.39	0.46	.50	.01
誤差	814.51	59	13.81			
指導法	0.45	1	0.45	0.02	.89	< .001
誤差	1444.05	59	24.48			

WLの優位性を示唆したWang(2023)は活動時間を制限しておらず、OLと比較してWLは約2倍の時間を要した。本研究では、両群の活動時間に差はなく、時間に制約を受けずに思考を深めることができるWLの利点が生かされず、即時的な言語化が求められる認知負荷の高いOLの学習効果が示されたのかもしれない。しかし、英文法テストの得点率は天井効果が見られた点や、自由英作文テストにおける正確性の向上には差が見られなかった点を踏まえると、形態の異なるランゲージングの学習効果を正確に把握するためには更なる検証が必要である。追加分析として、プレテスト(英文法)において正答率が8割未満だった参加者16名を対象に、プレテスト・ポストテスト(英文法・自由英作文)の得点を用いて2群間の差がないか統計分析を実施した。表5はプレテスト(英文法)において正答率が8割未満だった参加者16名のプレテスト・ポストテスト(英文法・自由英作文)の得点の記述統計を示す。

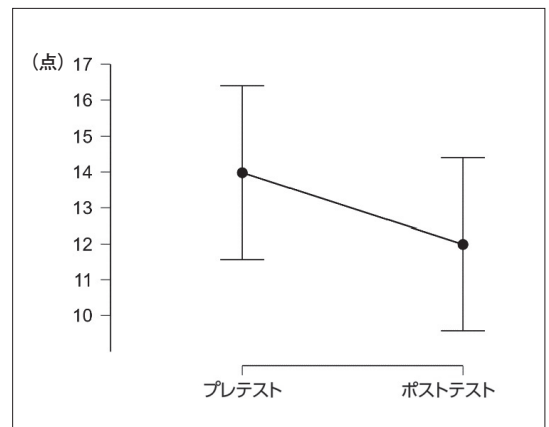
■表5: プレテスト・ポストテスト(英文法・自由英作文)の得点の記述統計($n = 16$)

英文法				自由英作文			
プレテスト		ポストテスト		プレテスト		ポストテスト	
平均値	標準偏差	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差
28.81	5.82	31.38	6.68	13.98	5.88	11.99	2.63

注. 英文法テストは43点満点、自由英作文の得点は「正規化した誤り数値(図2)」である。



■図5: 平均値のプロット(英文法)



■図6: 平均値のプロット(自由英作文)

データの正規性をシャピロ・ウィルク検定で確認したところ、英文法テストは認められたが、自由英作文テストは認められなかった。分析対象者が16名である点を鑑みて、プレテスト・ポストテスト(英文法・自由英作文)の得点の中央値を用いて、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実施した(表6)。

■表6: プレテスト・ポストテスト(英文法・自由英作文)得点のウィルコクソンの符号付順位和検定の結果($n = 16$)

	プレテスト		ポストテスト		z	p	r
	中央値	標準偏差	中央値	標準偏差			
英文法	30	33.00	33	6.68	-2.05	.04	.54
自由英作文	12.70	5.88	12.06	2.63	1.19	.25	.34

注. 英文法テストは43点満点、自由英作文の得点は「正規化した誤り数値(図2)」である。

結果、プレテスト・ポストテスト(英文法)の得点には有意な差が大程度の効果量(Cohen, 1988)と共に見られたが [$z(15) = -2.05, p = .04, r = .54$], プレテスト・ポストテスト(自由英作文)の得点には有意差が見られず、効果量も中程度(Cohen, 1988)だった [$z(15) = 1.19, p = .25, r = .34$]。つまり、プレテストにおいて比較的習熟度の低かった参加者は8回のランゲージング活動を経て英文法テストの得点が向上したことが示された。比較的習熟度の低い学習者に対しては、ランゲージングの形態に依らず学習効果が示唆された点は興味深い。つまり、習熟度の違いによって学習事項を「言語化する」行為を通して異なる学習効果が得られる可能性が示唆されたため、今後は個人差要因を含めてランゲージングの学習効果を検証することが期待される。

さらに、英文法テストにおいて8割未満の正答率であった関係詞[主格・目的格の who(m)/which/that] に焦点を当てた追加分析を行った。自由英作文テストのトピックは「尊敬する人」と「信頼する人」であり、関係詞を使用することがある程度期待されるだけでなく、1回目、2回目の処遇後に複数のLEを共有した際、両群共に関係詞に関する誤りとそのLE例を提示して指導を行ったため、より自由な言語産出における当該文法項目の正確性の向上を検証することは意義があると考えた。表7は、プレテスト・ポストテストで実施した自由英作文テストにおいて、参加者が産出した関係詞 who(m) と which, that の使用頻度、その中で正しく使用した回数と誤って使用した回数を示す。

■表7: 自由英作文テストにおける関係詞 who(m) と which, that の使用頻度 (N = 61)

		プレテスト			ポストテスト		
		頻度	正	誤	頻度	正	誤
who(m)	主格	75	69 (92)	6 (8)	102	62 (60.8)	40 (39.2)
	目的格	2	2 (100)	0 (0)	17	13 (76.5)	4 (23.5)
	計	77	71 (92.2)	6 (7.8)	119	75 (63)	44 (37)
which	主格	3	3 (100)	0 (0)	4	3 (75)	1 (25)
	目的格	1	1 (100)	0 (0)	0	0 (0)	0 (0)
	計	4	4 (100)	0 (0)	4	3 (75)	1 (25)
that	主格	3	1 (33.3)	2 (66.7)	1	0 (0)	1 (100)
	目的格	8	5 (62.5)	3 (37.5)	2	1 (50)	1 (50)
	計	11	6 (54.5)	5 (45.5)	3	1 (33.3)	2 (66.7)
合計		92	81 (88)	11 (12)	126	79 (62.7)	47 (37.3)

注. 「正」は正しい使用回数, 「誤」は誤った使用回数を示す。() 内は割合 (%) を示す。

関係詞 which と that の使用頻度は関係詞 who(m) と比較してとても少ない傾向が見られたが、やはり自由英作文のトピックが「尊敬する人」と「信頼する人」であった点が大きく影響したと思われる。ポストテストにおける関係詞 who(m) の使用頻度は、プレテストと比較して42回多く、参加者はより積極的に関係詞 who(m) を使用したことが伺える。そして使用頻度の増加に伴って、誤った使用回数が増加し、割合も高くなったことが読みとれる (7.8% → 37%)。つまり、自由英作文テストにおいて正確性が向上しなかった要因の一つとして、関係詞 who(m) の使用頻度が増加したことに伴う当該文法項目の誤用頻度の増加が考えられる。ランゲージング活動を継続することで、使用頻度が増加した関係詞 who(m) の使用の正確性が促進されるのか、更なる検証が必要である。

3.2 研究課題 2:LEは口頭と筆記で量的・質的に違いがあるか

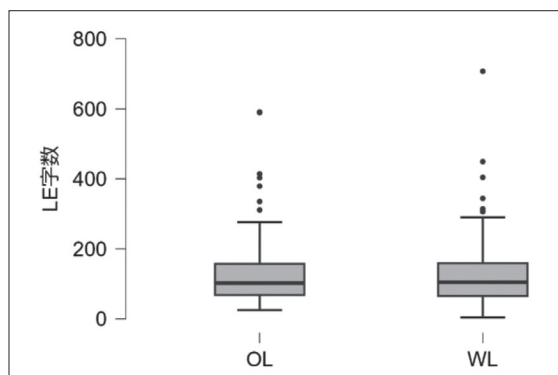
表8はOL群が記録したLE字数の記述統計、表9はWL群が記録したLE字数の記述統計を示す。記述統計の結果と図7の箱ひげ図からもわかるように、両群が記録したLE字数には量やばらつきの差がほとんど見られなかった。シャピロ・ウィルク検定で正規性が認められなかったため、両群が記録したLE字数の中央値を用いてノン・パラメトリック検定のマン・ホイットニー検定を行った結果、2群間に有意差は見られず効果量も非常に小さかった($U = 12322$, $p = .92$, $r = .01$)。参加者の多くは、中学で一人一台端末利用が促進されており、全ての参加者が高校入学時より半年間の授業内外で、一人一台端末を使用した課題や学習に取り組んでいるため、タイピング書字の経験がある程度豊富だったことが想定される。そのため、両群共に10分間与えられて実施したランゲージング活動においては、形態の違いはLE字数にほとんど影響を与えなかったと思われる。

■表8: OL群が記録したLE字数の記述統計

	課題①	課題②	課題③	課題④	課題⑤	課題⑥	課題⑦	課題⑧	全体
提出数	20	29	18	20	20	11	24	15	157
中央値	135	89	132	131	85	113	90	75	102
平均値	183.50	118.55	148.28	131.40	127.35	103.82	100.54	94.67	126.92
標準偏差	161.64	84.37	94.46	72.09	77.86	36.25	62.91	58.80	92.22
最小値	41	30	25	35	40	29	29	28	25
最大値	591	379	403	250	335	147	276	269	591

■表9: WL群が記録したLE字数の記述統計

	課題①	課題②	課題③	課題④	課題⑤	課題⑥	課題⑦	課題⑧	全体
提出数	19	21	21	19	18	19	22	19	158
中央値	120	118	98	103	92	90	129	104	105
平均値	147.58	162.05	122.38	108.90	102.78	115.11	126.05	122.90	126.53
標準偏差	149.58	102.44	86.96	66.67	61.78	70.54	83.83	88.35	92.40
最小値	14	38	10	22	13	32	4	14	4
最大値	707	449	344	236	223	270	290	404	707



■図7: OL群とWL群が記録したLE字数の箱ひげ図

表10は全体の課題①-⑧で与えられたIWCFの種類と、それに対するLEの種類を分類した頻度を示す。本研究において、参加者は合計761のIWCFが与えられて、その内515(67.7%)が文法に関するものであった。様々な習熟度を対象に異なるWCFを与えた多くの先行研究では、平均的に6-7割は文法面についてであることが示されているため(e.g., Fukuta et al., 2019; Suzuki et al., 2009; Suzuki, 2012; Yilmaz, 2016), 本研究においても類似の結果となった。また、生徒が記録したLE総数は636あり、与えられたIWCF数(761)に対して高い頻度(83.6%)でランゲージングに従事することが示された。さらに、記録されたLE総数636の内、414(65.1%)は最も学習効果が期待される「理由を伴った 正しい」レベルのLEであった。以上のように全体の傾向を把握した上で、口頭と筆記という形態の異なるLEは、5つの分類ごとの頻度に違いが見られるか検証した。

■表10: 課題①-⑧で与えられたIWCFの種類と、それに対するLEの種類 (全体)

	IWCF 分類				LE 分類					
	文法	語彙	他	合計	理正	理誤	単正	単誤	半信	合計
課題①	68	25	18	111	35	19	12	4	17	87
課題②	99	23	12	134	78	16	4	3	10	111
課題③	42	24	8	74	47	4	9	3	3	66
課題④	63	23	12	98	47	15	3	10	0	75
課題⑤	59	19	5	83	49	8	4	4	5	70
課題⑥	57	13	7	77	49	7	1	2	4	63
課題⑦	83	20	7	110	62	10	11	2	13	98
課題⑧	44	20	10	74	47	9	2	3	5	66
合計	515	167	79	761	414	88	46	31	57	636

注. 「理正」は「理由を伴った 正しい」, 「理誤」は「理由を伴った 誤った」, 「単正」は「単純な気づき 正しい」, 「単誤」は「単純な気づき 誤った」, 「半信」は「半信半疑」を示す(表1 参照)。

表11はOL群の課題①-⑧で与えられたIWCFの種類とそれに対するLEの種類を分類した頻度を示す。生徒が記録したLE総数は306あり、与えられたIWCF数413に対して全体平均(83.6%)より低い率(74.1%)でランゲージングに従事したことが示された。さらに、記録されたLE総数306の内、182(59.5%)は最も学習効果が期待される「理由を伴った 正しい」レベルのLEであった。一方、表12が示す通り、WL群が記録したLE総数は330あり、与えられたIWCF数348に対して全体平均(83.6%)より高い率(94.8%)でランゲージングに従事したことが示された。さらに、記録されたLE総数330の内、232(70.3%)は最も学習効果が期待される「理由を伴った 正しい」レベルのLEであった。つまり、WL群はOL群と比較して、与えられたIWCFに対するランゲージング活動により高い関与を示し、記録されたLEの質は気づきレベルが最も高い「理由を伴った 正しい」レベルになる傾向だったと言える。

■表11: 課題①-⑧で与えられたIWCFの種類とそれに対するLEの種類 (OL群)

	IWCF 分類				LE 分類					
	文法	語彙	他	合計	理正	理誤	単正	単誤	半信	合計
課題①	46	14	8	68	15	10	8	3	8	44
課題②	64	8	7	79	39	8	4	3	7	61
課題③	20	16	5	41	22	2	5	2	2	33
課題④	31	12	8	51	25	7	2	2	0	36
課題⑤	33	13	3	49	22	5	4	3	3	37
課題⑥	29	3	3	35	16	5	1	0	2	24
課題⑦	43	12	5	60	28	3	7	1	9	48
課題⑧	12	11	7	30	15	6	0	0	2	23
合計	278	89	46	413	182	46	31	14	33	306

注. 「理正」は「理由を伴った 正しい」, 「理誤」は「理由を伴った 誤った」, 「単正」は「単純な気づき 正しい」, 「単誤」は「単純な気づき 誤った」, 「半信」は「半信半疑」を示す(表1参照)。

■表12: 課題①-⑧で与えられたIWCFの種類とそれに対するLEの種類 (WL群)

	IWCF 分類				LE 分類					
	文法	語彙	他	合計	理正	理誤	単正	単誤	半信	合計
課題①	22	11	10	43	20	9	4	1	9	43
課題②	35	15	5	55	39	8	0	0	3	50
課題③	22	8	3	33	25	2	4	1	1	33
課題④	32	11	4	47	22	8	1	8	0	39
課題⑤	26	6	2	34	27	3	0	1	2	33
課題⑥	28	10	4	42	33	2	0	2	2	39
課題⑦	40	8	2	50	34	7	4	1	4	50
課題⑧	32	9	3	44	32	3	2	3	3	43
合計	237	78	33	348	232	42	15	17	24	330

注. 「理正」は「理由を伴った 正しい」, 「理誤」は「理由を伴った 誤った」, 「単正」は「単純な気づき 正しい」, 「単誤」は「単純な気づき 誤った」, 「半信」は「半信半疑」を示す(表1参照)。

次に, OL 群と WL 群が記録した LE の5つの分類ごとの頻度差を χ^2 検定で検証した。結果, 表13が示すとおり, 両群間には小程度の効果量 (Cohen, 1988) と共に有意差が見られ [$\chi^2(4) = 12.99, p < .05, V = .13$], 残差分析と多重比較の結果, WL 群が記録した「理由を伴った 正しい」LE 頻度は有意に多く, 「単純な気づき 正しい」LE 頻度は有意に少ないことが示された。さらに, OL 群が記録した「理由を伴った 正しい」LE 頻度は有意に少なく, 「単純な気づき 正しい」LE 頻度は有意に多いことが示されたため, 形態の異なるランゲージングによる LE の質は異なる可能性が示唆された。つまり, WL 群の LE は「理由を伴った 正しい」LE になる傾向が示されただけでなく, 「単純な気づき 正しい」LE にならない傾向が見られ, より質の高い LE になる可能性が示された。一方, OL 群の LE は「理由を伴った 正しい」LE にならない傾向が示されただけでなく, 「単純な気づき 正しい」LE になる傾向が見られ, 比較的質の低い LE になる可能性が示唆された。

■表13: OL群とWL群が記録したLEの分類別頻度数のクロス表

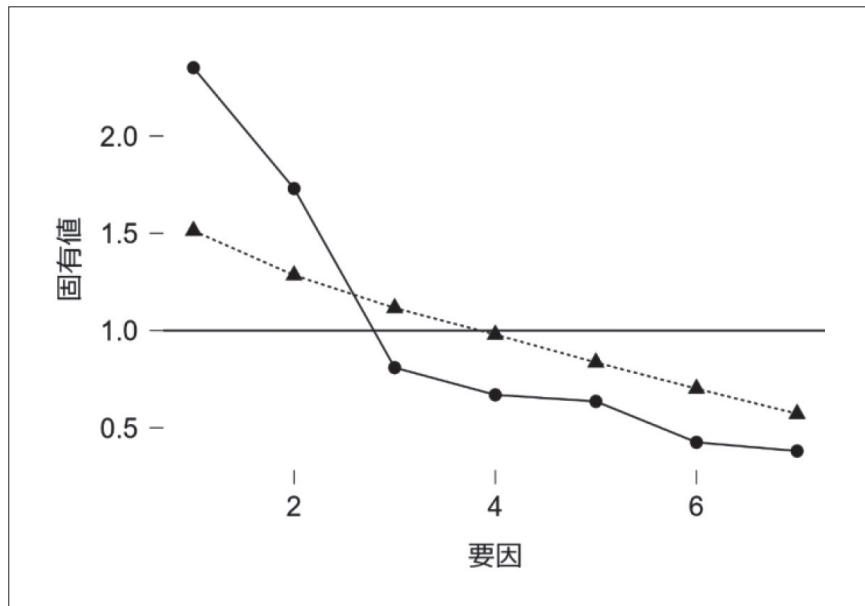
		LE 分類					合計	
		理 正	理 誤	単 正	単 誤	半 信		
形態	WL	度数	232	42	15	17	24	330
		調整済み残差	2.9	−0.8	−2.7	0.3	−1.5	
		p 値	<.01	ns	<.01	ns	ns	
	OL	度数	182	46	31	14	33	306
		調整済み残差	−2.9	0.8	2.7	−0.3	1.5	
		p 値	<.01	ns	<.01	ns	ns	
合計		度数	414	88	46	31	57	636

注. ns = non-significant[有意差なし]を示す。「理正」は「理由を伴った 正しい」,「理誤」は「理由を伴った 誤った」,「単正」は「単純な気づき 正しい」,「単誤」は「単純な気づき 誤った」,「半信」は「半信半疑」を示す(表1参照)。

本研究の結果より,同じ時間が与えられて個人で取り組むランゲーjingの場合,OLよりもWLで記録されるLEの質は高くなることが示唆された。Peng et al.(2023)では,個人で取り組むWL群はペアで取り組むOL群よりも,WCFに対するランゲーjingに従事する傾向にあり,かつ書き直しにおける高い修正率を報告している一方で,個人では解決できないWCFに対しては認知処理レベルの低いLEが多いことも明らかになっている。つまり,個人で取り組むWLはランゲーjing活動に関与する度合いは強いが,学習者が理解できる範囲を超えた問題の場合は質の低いLEをもたらすことが予想される。本研究では,与えられたフィードバックの明示性が参加者にとってある程度適切だったため,気づきレベルの高い「理由を伴った 正しい」LEが多く記録されたと考えられる。さらに,個人で取り組むWLは,筆記形式故に学習者はより深い思考を伴ったため,LEの質の観点でOLよりも優位性が示されたと考えられる。また,本研究におけるOLの様相も学習効果やLEの質に影響した可能性がある。つまり,OL群は思考した内容を即自的に言語化するのではなく,ある程度思考を整理した後に口頭で言語化した内容を録音して提出する方法だった。課題提出率がWL群と比較して20%以上低かったため,OL群はランゲーjing課題に対する関与が低かったことでLEの質が上がりなかった可能性や,誤りの原因について思考はしたが課題を提出しなかった可能性が考えられる。いずれにせよ,Peng et al.(2023)の結果と類似して,気づきレベルの高いLEが必ずしも高い学習効果につながらない可能性が示されたため,LEの質と学習効果の関係を明らかにするために,ターゲットとなる英文法を設定する等,より焦点を絞った検証が必要と考える。

3.3 研究課題3: [研究課題2]で違いが認められた場合,英語ライティングに対する考えとLEの頻度や質はどのような関係があるか

英語ライティング課題が返却された後の学習行動や,学習時の「言語化」に関する考えについて問う英語ライティングに関する質問紙調査を実施した(資料2)。回答は,1:「まったくない」から6:「とてもよくある」,1:「まったくない」から6:「すごくある」,あるいは,1:「まったくない思わない」から6:「とてもよく思う」のいずれかの自己評価である。全質問項目のクロンバックの信頼性係数は $\alpha = .70$ であり最低限望ましい信頼性を示した。スクリー基準をもとに固有値の落差と値,および因子の項目内容を考慮した上で,2因子解を想定して因子分析(最尤法・オブリンミン回転)を行った(図8)。表14のとおり,因子1は項目4,6,3,7が高い因子負荷を示したため,「英語ライティング返却後の学習行動」因子と命名した。因子2は項目15,14,16が高い因子負荷を示したため,「学習時の『言語化』に対する趣向」因子と命名した。なお,因子間相関は因子1と因子2の間は $r = .11$,累積因子寄与率は.45であった。



■図8: スクリープロット

■表14: 各因子の下位項目の記述統計と探索的因子分析の結果

項目	記述統計				パターン行列	
	平均値	標準偏差	尖度	歪度	因子1	因子2
4	5.70	0.56	-1.77	2.29	.87	-.03
6	5.43	0.83	-1.32	0.91	.65	.03
3	5.46	0.87	-2.64	10.6	.57	.01
7	4.46	1.04	-0.39	0.66	.37	.17
15	5.25	0.87	-1.30	2.17	.02	.89
14	4.67	1.14	-1.15	1.63	-.17	.59
16	4.38	0.88	-0.22	-0.09	.17	.50
因子間相関					因子1	因子2
因子1					—	
因子2					.11	—
クロンバックα					.67	.66

表15は因子1「英語ライティング返却後の学習行動」と因子2「学習時の『言語化』に対する趣向」、両群のLEの記述統計を示す。表16は参加者の英語ライティング課題が返却された後の学習行動と形態の異なるLEの頻度と質の関係を検証するため、探索的因子分析の結果得られた因子1と因子2の平均値と記録されたLE従事率(%), LE総数, 「理由を伴った 正しい」LE数, LE総字数を用いて相関分析を行った結果である。なお、シャピロ・ウィルク検定で正規性が認められなかったため($p < .001$)、スピアマンの順位相関分析を行った。

■表15: 因子1と因子2, LEの記述統計 (N = 61)

		因子1	因子2	LE従事率	LE総数	理正LE数	LE総字数
WL (n = 27)	平均値	5.01	4.79	95.59	12.26	8.59	693.63
	標準偏差	0.59	0.61	9.03	5.80	4.06	396.36
	中央値	5.25	5.00	100	11	8	663
OL (n = 34)	平均値	5.41	4.75	74.13	9.00	5.35	583.59
	標準偏差	0.57	0.85	28.46	6.50	3.89	458.04
	中央値	5.50	5.00	82.84	9	5	474

注. 「LE従事率」は「LE総数 / IWCF総数」を示す。「理正LE数」は「理由を伴った正しい」LE数を示す。

■表16: スピアマンの順位相関行列 (WL群/OL群)

	因子1	因子2	LE従事率	LE総数	理正LE数	LE総字数
因子1	—					
因子2	.27 / .20	—				
LE従事率	.26 / -.20	.11 / .25	—			
LE総数	-.11 / -.26	-.09 / .42*	-.06 / .44**	—		
理正LE数	.01 / -.16	-.03 / .28	.09 / .51**	.87*** / .79***	—	
LE総字数	-.28 / .22	-.44* / .35*	-.04 / .49**	.50** / .84***	.54** / .81***	—

注. *. 相関係数は5%水準で有意(両側), **. 相関係数は1%水準で有意(両側), ***. 相関係数は0.1%水準で有意(両側)

結果, OL群のLE従事率と因子1の間に中程度の効果量(Cohen, 1988)と共に弱い負の相関関係が見られたが, WL群のLE従事率と因子1の間は中程度の効果量(Cohen, 1988)と共に弱い正の相関関係だった(それぞれ $\rho = -.20, p > .05$; $\rho = .26, p > .05$)。因子1は「英語ライティング返却後の学習活動」因子であり, ライティング課題が返却された後に学習活動を行う傾向のある学習者ほど, OLへの関与の度合いが低く, 対してWLにはより取り組んだことが示された。また, OL群のLE従事率と因子2の間に中程度の効果量(Cohen, 1988)と共に弱い正の相関関係が見られたが, WL群のLE従事率と因子2の間は非常に弱い正の相関関係だった(それぞれ $\rho = .25, p > .05$; $\rho = .11, p > .05$)。因子2は「学習時の『言語化』に対する趣向」因子であり, ランゲージングに対する肯定的な印象を持つ学習者ほど, OLへの関与の度合いが強いが, WLはその趣向の程度との関連が弱い可能性が示唆された。

また, OL群のLE総数と因子1の間に中程度の効果量(Cohen, 1988)と共に弱い負の相関関係が見られたが, WL群のLE総数と因子1の間には非常に弱い負の相関関係だった(それぞれ $\rho = -.26, p > .05$; $\rho = -.11, p > .05$)。因子1が含む事後活動は因子2が含む学習事項の「言語化」以外の行動であり, 学習者が英語ライティング課題の返却後に「言語化」以外の学習活動を行う傾向にある場合, OLはIWCFに対するLE数が減少することが示された。一方, WLは, 「言語化」以外の学習活動をする傾向の学習者であっても, LE総数が減少する程度は小さいと言える。また, OL群のLE総数と因子2の間に中程度の効果量(Cohen, 1988)と共に中程度の正の相関関係が見られたが, WL群のLE総数と因子2の間は非常に弱い負の相関関係だった(それぞれ $\rho = .42, p < .05$; $\rho = -.09, p > .05$)。LE従事率と同様に, ランゲージングに対する肯定的な印象を持つ学習者ほどOLではより多くのLE数を記録し, WLはその趣向の程度と関連は弱い可能性が示された。

さらに、OL 群の LE 総字数と因子2の間には弱い正の相関関係が見られたが、WL 群の LE 総字数と因子2の間に中程度の効果量 (Cohen, 1988) と共に中程度の負の相関関係が見られた (それぞれ $p = .35, p > .05$; $p = -.44, p < .05$)。同様の傾向は、OL 群と WL 群それぞれの LE 総字数と因子1の間にも見られた (それぞれ $p = .22, p > .05$; $p = -.28, p > .05$)。つまり、ライティング課題が返却された後に学習活動を行う傾向のある学習者は、WL では記録される LE 字数が減少する傾向があり、特にランゲージングに対する肯定的な印象を持つ学習者ほどその関連が強かった。学習者自身のペースでゆっくり取り組める WL は深い思考を促すと言われているが、その形態の特徴が思考を整理することを容易にし、結果として記述量の少ないまとまった説明がなされるのかもしれない。また、OL 群の LE 総字数と因子2の間に見られる弱い正の相関関係は、OL に取り組む場合、ランゲージングに対する肯定的な印象を持つ学習者ほどより多くの説明をする傾向を示している。

さらに興味深い結果として、LE 従事率とそれぞれの LE 総数、「理由を伴った 正しい」LE 数、LE 総字数の間に見られる関係は OL 群と WL 群では大きな違いが見られた。OL 群の LE 従事率と LE 総数の間に中程度の効果量 (Cohen, 1988) と共に中程度の正の相関関係が見られたが、WL 群の LE 従事率と LE 総数の間は非常に弱い負の相関関係だった (それぞれ $p = .44, p < .01$; $p = -.06, p > .05$)。OL 群の LE 従事率と「理由を伴った 正しい」LE 数の間に効果量大 (Cohen, 1988) と共に中程度の正の相関関係が見られたが、WL 群の LE 従事率と「理由を伴った 正しい」LE 数の間は非常に弱い正の相関関係だった (それぞれ $p = .51, p < .01$; $p = .09, p > .05$)。OL 群の LE 従事率と LE 総字数の間に中程度の効果量 (Cohen, 1988) と共に中程度の正の相関関係が見られたが、WL 群の LE 従事率と LE 総字数の間は非常に弱い負の相関関係だった (それぞれ $p = .49, p < .01$; $p = -.04, p > .05$)。つまり、OL に取り組む場合は、ランゲージング活動に取り組む頻度と記録される LE 字数が多いほど LE 従事率も高くなり、かつ記述される内容が理由を伴った正しい LE になる傾向が示された。対して、WL ではそのような関連が弱く、OL と比較して学習者の個人差要因が小さい形態のランゲージングである可能性が示唆された。

形態の異なるランゲージングと記録される LE の量・質的データの関係性を踏まえると、効果的なランゲージング活動を目指して OL を実践する場合、WL よりも多くの指導や介入が必要だと考えられる。つまり、英語ライティング課題の返却後に行う学習活動を促すだけでなく、言語面の誤り等に対して理由を考えて「言語化」するランゲージングの学習効果を伝えることで、生徒は「言語化する」学習方法に対してより肯定的な印象を持つようになり、結果として OL により従事すると考えられる。

質問紙調査を実施した約1ヶ月後に、英語ライティングに関する質問紙の回答結果より、因子1と因子2の値が共に平均値より0.5ポイント以上高かった2名の参加者を対象に、今回実施したランゲージング活動に関するインタビュー調査を行った (20-30分程度)。参加者の発言から、本研究で継続して実施した「自由英作文課題+ランゲージング活動」を通して、英語ライティング課題に取り組む際の態度面の変化があった可能性が示唆された (表17)。今後のランゲージング研究は、言語学習における学習効果のみを対象とするのではなく、学習者の英語ライティングに対する態度面や学習方略への影響についても検証することで、より多くの教育的示唆が得られると思われる。

■表17: ランゲージング活動に関する参加者の発言内容(筆者による書き起こし)

参加者 A	単純なミスはすく目につきやすくなりました。三単現の漏れとか…そういうのが、自分でもそういうのすら、なんでしょう。どこまで大幅な間違いをしたんだと思ってずっと見てたのが、その程度のものは、もうなんでそのミスしてるんだろうっていうのを数えられるようになりました。
参加者 B	なんか最初は結構、まあ正直あまり、こうバツとかもらいたくない、みたいな気持ちが大きくて、こう、知ってる、すごい中学生文法みたいなので全部書いてたりしたんですけど、後半になるにつれて、逆にもらったらこれ使い方違うんだってわかるからいいのかもしれないと思って、ちょっと難しい辞書で引いた使い方を当てはめて使ってみて、みたいなことは、をするようになったのはちょっと変化かなって思います。

4 まとめ

4.1 本研究の結論

本研究は、第二言語習得研究において言語学習を促進することが示唆されているランゲージングの学習効果と記録されるLEの頻度や質を、口頭と筆記という異なる形態に注目して検証した。日本語を母語とする高校生を対象に、5ヶ月間にわたって自由英作文課題＋ランゲージング活動を8回行った。

【研究課題 1】「長期的にOLまたはWLに取り組むことで言語面の正確性は向上するか」に対しては、英文法テストの得点はOL群が有意に向上したことが示された一方、WL群は変化が見られなかった。自由英作文テストにおける正確性は両群ともに変化が見られなかったため、より自由な言語産出の場合にランゲージングの学習効果は両群共に見られなかったと言える。英文法テストの得点は天井効果を示したため、プレテストにおいて8割未満の正答率だった16名を対象に追加分析した結果、プレーポスト間で有意に向上したことが示されたため、比較的習熟度が低い学習者はよりランゲージング活動の恩恵を受けた可能性が示唆された。また、英文法テストにおいて比較的正答率が低かった関係詞who(m)とwhich, thatの自由英作文テストにおける使用頻度、その内正しい使用回数および誤った使用回数を分析した結果、参加者は関係詞who(m)をより積極的に使用することに伴って誤用頻度も増加していたことが明らかになった。

【研究課題 2】「LEは口頭と筆記で量的・質的に違いがあるか」に対しては、同じ活動時間の場合、記録されるLE字数には両群間で差が見られなかった。一方、LEの質については、WL群がOL群と比較してより高い頻度でランゲージングに取り組み、かつ気づきレベルの高い「理由を伴った 正しい」LEをより多く産出したことが明らかになった。

【研究課題 3】「【研究課題 2】で違いが認められた場合、英語ライティングに対する考えとLEの頻度や質はどのような関係があるか」に対しては、OLの場合、ランゲージングに対する肯定的な印象を持つ学習者ほど、よりランゲージング活動に従事して多くのLE数やLE字数を記録したが、WLの場合はランゲージングに対する趣向の程度は、活動への関与の度合いや記録されるLE数と関連が弱いことが示された。さらに、OLの場合のみ、ランゲージング活動に取り組む頻度や記録されるLE字数が多いほどLE従事率は高くなり、かつ記述される内容が気づきレベルの高いLEになる傾向が見られた。つまり、OLとWLは学習者の個人差要因が与える影響の度合いが異なる活動である可能性が示唆された。

本研究の結果が示唆する興味深い内容を2点述べる。まず、個人で取り組むランゲージングの学習効果や記録されたLEを量的・質的に比較検証した結果、口頭と筆記の形態の大きな差はないことが示唆された。教室現場の状況や活動の目標に合わせてOLとWLを選択して実施したり、両方を活用したりすることも可能だと考えられる。次に、同じ制限時間が与えられて個人でランゲージングに取り組む場合、LE総数や従事率はWLの方がOLよりも高かったが、OLのみ、学習時の「言語化」に対する肯定的な印象の程度がランゲージング活動への関与の度合いや記録されるLEの量や質に強く関連することが示唆された。学習者自身が「与えられたフィードバックについて理由を考えて言語化すること」は効果的な学習方法だと認識することで、ライティング課題の返却後に行うランゲージング活動は、さらに効果的な学習方法になると思われる。

4.2 課題と展望

本研究は、教室現場における長期的なランゲージング活動の学習効果とLEの内容を検証した点で意義があるが、研究デザイン上の課題が多かったことも認識している。以下、課題を4点述べると共に、どのように今後の英語指導や英語学習に生かせるか展望を述べる。

1点目は、OL群とWL群が取り組んだランゲージング活動の高い類似性である。本研究では、両群共に10分間でGoogle フォームの回答を送信するよう指示され、OL群は口頭で説明する内容を録音し、WL群

はタイピング書字で記述して提出した。つまり、OL群も即時的な言語化は求められておらず、ある程度思考を整理した後でOLを行った点において、OLの特徴がなくなり、結果としてWLとほとんど違いのない学習活動だったと言える。結果的に、学習者に考える時間を十分に与えることでOLとWLのどちらの形態であっても同等の学習効果が得られることが示唆された。形態の異なるランゲーjingの学習効果をより正確に比較検証するために、今後はより短い時間を設定してOL群に即時的な言語化を促した場合の検証も求められる。

2点目は、タイピング書字のWL群を設定した点である。一人一台の端末利用が促進されている教室現場では、学習の振り返り活動を手書きではなくタイピング書字で行うことは一般的になりつつあるため、ランゲーjing活動をタイピング書字で行うことは自然な印象を受けるかもしれない。しかし、タイピング書字は経験の差が記録される情報量に密接に関わる点や、書字活動や学習活動の種類によっては、手書きとは異なる学習効果をもたらす可能性が先行研究で示唆されている(e.g., Bui et al. 2013; 川原, 2020)。つまり、英語学習におけるランゲーjing活動は手書きとタイピング書字でどのような学習効果の違いがあるか、記録されるLEの量や質に違いがあるかは明らかになっていない。本研究の結果は、WL群が記録したLE字数はOL群と差がなく、むしろ質的には優位性が見られたものの、より高い学習効果は示されなかった。Wang (2023) や Peng et al. (2023) は手書きのWL群を設定しており、本研究ではタイピング書字のWL群だった点が結果に影響を与えた可能性が考えられる。つまり、手書きの場合、記録される情報量は減ることが予想されるが、学習者は思考を整理してまとめることが求められるため、より気づきレベルの高いLEが多く記録されて言語学習を促進するかもしれない。

3点目は、ライティング力に対する正確性の分析方法の妥当性である。本研究における言語面の正確性は、記述式の空所補充問題である英文法テストの得点と、自由英作文テストにおける正規化した誤り数値で検証した。まず、英文法テストは天井効果が見られたため、言語面の正確性が向上したか否かを分析する上で妥当性が低かった。また、8回のランゲーjing活動では特定のターゲットとなる英文法を設定しておらず、英語ライティングにおける正確性全般の学習効果を検証したため、処遇の学習効果を詳細に解釈することができなかった。ランゲーjingの先行研究では、処遇が効果的な文法項目とそうでないものがあることが示唆されている(石川・鈴木, 2022)。その点を踏まえると、高等学校で初めて扱う仮定法過去完了をターゲットの英文法に設定して、ランゲーjing活動が既習事項の仮定法過去と比較して新出文法事項の仮定法過去完了の理解度をどれだけ促進するかを比較検証する等の研究は意義が大きいだろう。また、明示的知識が少ない習熟度の学習者や中学生を対象とした場合、形態の異なるランゲーjingの学習効果は本研究の結果と異なるのかについても検証することが必要と考える。

4点目は、英語ライティングに関する質問紙項目の妥当性と信頼性である。本研究では、英語ライティング課題が返却された後の学習活動や学習時の「言語化」に関する考えについて問う英語ライティングに関する質問紙調査を実施した(資料2)。全質問項目や因子1と因子2のクロンバックの信頼性係数は決して高くなく(それぞれ $\alpha = .70, .67, .66$)、質問紙尺度として信頼性が不十分であった可能性が高い。学習者のライティング課題返却時の学習活動とランゲーjing活動に対する考えの関係を明らかにするために、より高い信頼性が十分に認められる質問紙尺度を作成する必要がある。

ランゲーjingは「自律的な学習者」の育成にも寄与する可能性を踏まえると、個人でランゲーjingに取り組む意義は大きいだろう。しかし、学習者の習熟度によって個人では解決できない問題に直面する場合も多く、ペアやグループで取り組む協働的なランゲーjing活動も指導法の選択肢に入る(e.g., Storch, 2011; Swain et al., 2013)。特にライティング指導の文脈では、学習者は目的・場面・状況に応じた適切な言語使用のみでなく、言語面の正確性にも注意しながら思考し改訂する「自立した書き手(広瀬, 2014, p.132)」を目指す上で、様々な形態のランゲーjing活動が適切な支援になり得ると考える。

謝辞

本研究を実施する貴重な機会を与えて下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御礼申し上げます。特に、本研究の実施・報告にあたり、研究助言者の西垣知佳子先生には貴重なご助言ご指導を数多く賜りました。また、筑波大学の名畑目真吾先生にも多くの有益なご助言を賜りました。この場をお借りして、心より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

- ベネッセ教育総合研究所 (2014). 「中高生の英語学習に関する実態調査2014」
- Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology*, 105(2), 299-309. <https://doi.org/10.1037/a0030367>
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.
- 江下陣 (2020). 「学習者のスピーチとモデルスピーチの比較による主体的な『気づき』を促すスピーキング指導」『第32回「英検」研究助成報告」STEP BULLETIN. Vol. 32, 108-124.
- Fukuta, J, Tamura, Y, & Kawaguchi, Y. (2019). Written languaging with indirect feedback in writing revision: Is feedback always effective? *Language Awareness*, 28(1), 1-14. <https://doi.org/10.1080/09658416.2019.1567742>
- 広瀬恵子 (2014). 「第5章 ライティング 3 ライティングのフィードバック」全国英語教育学会 (編). 『英語教育学の今: 理論と実践の統合』(pp. 132-135) 全国英語教育学会.
- 市川裕理 (2022). 「ランゲージングにおける『協働的足場かけ』の研究」『中部地区英語教育研究紀要』50, 17-24.
- 今澤真紀 (2024, 6月1日, 於筑波大学). 「日本人EFL学習者のライティングにおける協同的ランゲージング (Collaborative Languaging) の効果」第3回言語学習評価研究会発表資料.
- 石川正子・鈴木渉 (2022). 「筆記ランゲージングを取り入れた教室実践の試み」『城西大学語学教育センター研究年報』14, 1-11. https://doi.org/10.20566/18801919_14_1
- Ishikawa, M. (2015). Metanotes (written languaging) in a translation task: Do L2 proficiency and task outcome matter? *Innovation in Language Learning and Teaching*, 9(2), 115-129. <https://doi.org/10.1080/17501229.2013.857342>
- Kazemi, P., Pourdana, N., Khalili, G. F., & Nour, P. (2022). Microgenetic analysis of written languaging attributes on form-focused and content-focused e-collaborative writing tasks in Google Docs. *Education and Information Technologies*, 27(8), 1-24. <https://doi.org/10.1007/s10639-022-11039-y>
- 川原淳 (2020). 「『手書き』と『タイピング書字』における脳の言語機能に及ぼす影響の比較—NIRSを用いた統計的な調査—」『早稲田大学国語教育研究』40, 63-73.
- 国立教育政策研究所 (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 小学校外国語・外国語活動」
- 駒井健吾 (2024). 「生徒の『思い』を伸ばすライティング指導と評価の工夫」『英語教育2024年6月号』20-21. 大修館書店.
- Li, C., Liu, B., & Wang, X. (2023). A scoping review of research on languaging in second language education. *Heliyon*, 9(5), 1-12. <https://doi.org/10.1016/j.heliyon.2023.e16300>
- 松下信之 (2024). 「無理なく効果的に行えるライティングのフィードバック」『英語教育2024年6月号』34-35. 大修館書店.
- 文部科学省 (2018). 「平成29年度英語力調査結果 (高校3年生) の概要」
- 西垣知佳子・川名隆行・見目慎也・中井康平・山崎達也 (2024). 「中学生が身に付けている英文法知識の調査」『関東甲信越英語教育学会誌』38, 127-140.
- Niu, R. & Li, L. (2017). A review of studies on languaging and second language learning (2006-2017). *Theory and Practice in Language Studies*, 7(12), 12, 1222-1228. <http://dx.doi.org/10.17507/tpls.0712.08>
- 布川裕行 (2018). 「ランゲージングのL2文法習得に対する効果—書記フィードバックにおける促進的役割—」『東北大学高度教養教育・学生支援機構紀要』4, 339-353.
- Peng, C. X., Storch, N., & Knoch, U. (2023). Greater coverage vs. deeper processing? Comparing individual and collaborative processing of teacher feedback. *Language Teaching Research*, 1-27. <https://doi.org/10.1177/13621688231214910>
- 斎藤玲・鈴木渉 (2022). 「第二言語ライティングにおける書記訂正フィードバックとその振り返り, 書き直しに関する態度と行動: 認知モデルの提案と予備的調査の報告」『東北英語教育学会研究紀要』42, 49-64.
- 佐藤貴弘 (2020). 「協同学習が中学校初級レベル英語学習者の第二言語習得に与える影響」『中部地区英語教育学会研究紀要』49, 141-148.
- Storch, N. (2011). Collaborative writing in L2 contexts: Processes, outcomes, and future directions. *Annual review of applied linguistics*, 31, 275-288. <https://doi.org/10.1017/S0267190511000079>
- 鈴木渉 (2008). 「ランゲージングが第二言語学習に与える効果」『第20回「英検」研究助成報告」STEP BULLETIN. Vol. 20, 60-75.
- Suzuki, W. (2012). Writing languaging, direct correction, and second language writing revision. *Language Learning*, 62(4), 1110-1133. <https://doi.org/10.1111/j.1467-9922.2012.00720.x>
- Suzuki, W. & Itagaki, N. (2009). Languaging in grammar exercises by Japanese EFL learners of differing proficiency. *System*, 37(2), 217-225. <https://doi.org/10.1016/j.system.2008.10.001>
- Suzuki, W. & Storch, N. (2020). *Languaging in Language Learning and Teaching: A Collection of Empirical Studies (Language Learning & Language Teaching)*. Amsterdam: John Benjamins.

引用文献

- Swain, M. (2006). Linguaging, agency and collaboration in advanced second languaging proficiency. In H. Byrnes (Ed.), *Advanced Languaging Learning: The contribution of Halliday and Vygotsky* (pp. 95-108). Continuum.
- Swain, M., & Lapkin, S. (2013). Focus on form through collaborative dialogue: Exploring task effects. In *Researching pedagogic tasks* (pp. 99-118). Routledge.
- Swain, M., & Watanabe, Y. (2013). Linguaging: Collaborative dialogue as a source of second language learning. *The encyclopedia of applied linguistics*, 3218-3225. <https://doi.org/10.1002/9781405198431.wbeal0664.pub2>
- 高木哲也 (2022a). 「高校英語授業における間接書記訂正フィードバックに対する筆記ランゲージングの効果」『関東甲信越英語教育学会誌』36, 141-154. https://doi.org/10.20806/katejournal.36.0_141
- 高木哲也 (2022b). 「ライティング指導の一環で筆記ランゲージングを試みる—筆記ランゲージング・エピソードの分析に関する一考察—」『英語授業研究学会紀要』31, 98-109.
- 高木哲也 (2023a). 「『時制』の習熟を目指したディクトグロスの実践—ICTを活用した筆記ランゲージングを加えて—」『筑波大学附属高等学校研究紀要』64, 65-77.
- 高木哲也 (2023b). 「筆記ランゲージング研究の動向と展望—日本の英語授業における活用の可能性と課題—」『関東甲信越英語教育学会誌』37, 183-196. https://doi.org/10.20806/katejournal.37.0_183
- 高木哲也 (2024). 「高校生が繰り返し明示的知識を外化する実践—筆記ランゲージング・エピソードの質的分析—」『英語授業研究学会紀要』32, 71-83.
- Ushiyama, M. (2022). The Effect of Linguaging on the Grammatical Accuracy of Writing and Speaking English as a Foreign Language in Junior High School. Doctor of Philosophy (Humanities) Tokyo University of Foreign Studies, JAPAN.
- Wang, W. (2023). The efficacy of written corrective feedback: Searching for the best predictors. *International Journal of Applied Linguistics*, 34(2), 1-17. <https://doi.org/10.1111/ijal.12503>
- Yilmaz, M. (2016). Improving Turkish EFL learners writing accuracy: Effects of written languaging and languaging type. *Procedia - Social and Behavioral Sciences*, 232, 413-420. <https://doi.org/10.1016/j.sbspro.2016.10.057>
- 吉川庸子 (2024). 「振り返りを活用して児童の言語への気付きを促す—小学校英語教育におけるランゲージング」『英語教育2024年1月号』40-41. 大修館書店.

資料1: Google フォームを使用した英文法テスト (プレテスト冒頭部分)

45 セクション中 2 個目のセクション

無題のセクション

日本語の文に合うように、() に入る英語をキーボードから入力しましょう。

- 全部で43問あります。
- 制限時間は20分です。
- 回答する時は、回答欄の一番左から入力を始めます。
- 回答が1つの単語の時 単語の前後にスペースを入れないように。
- 回答が2つの単語の時 単語と単語の間に1文字分のスペースを入れます。その他のスペースを入れないように。

セクション 2 以降 次のセクションに進む

45 セクション中 3 個目のセクション

セクション タイトル (省略可)

説明 (省略可)

問 1 *

それをしないで。
() do it

記述式テキスト (短文回答)

資料2: 英語ライティングに関する質問紙項目と回答方法

セクション	質問項目	回答方法
英語ライティング全般について	1. 英語ライティングに対する自信の度合いを教えてください。	6件法 (1: まったくない, 2: ない, 3: あまりない, 4: 少しある, 5: ある, 6: すごくある)
	2. 英語ライティングの力を高めたいと思いますか。	
英語ライティングで教師からフィードバックを受けた後の振り返りにについて	3. 英語ライティングが教員から返却されたとき, 得点や評価を確認する。	6件法 (1: まったくない, ー6: とてもよくある)
	4. 英語ライティングが教員から返却されたとき, 訂正されたところを確認する。	
	5. 英語ライティングが教員から返却されたとき, 訂正の有無にかかわらず全体を確認する。	
	6. 英語ライティングが教員から返却されたとき, 訂正された理由や原因を考える。	
	7. 英語ライティングが教員から返却されたとき, 自分の苦手なポイントやクセについて考える。	
	8. 英語ライティングが教員から返却されたとき, 訂正された箇所そのものについて参考書等で調べる。	
	9. 英語ライティングが教員から返却されたとき, 訂正された箇所と関係する・関連する内容について参考書等で調べる。	
	10. 英語ライティングが教員から返却されたとき, 訂正された理由や原因をAI添削ツールで調べる。	
	11. 英語ライティングが教員から返却されたとき, 訂正された理由や原因を考えて言語化する(口頭で説明する,または書き記す)。	
教師からのフィードバックと「ライティング振り返り活動」そのものについて	12. 英語ライティングにおいて, 教師からのフィードバックは効果的な学習だと思いますか。	6件法 (1: まったく思わない, ー6: とてもよくそう思う)
	13. 英語ライティングにおいて, 教師からのフィードバックに加えて「誤りの理由を考えて言語化すること」は効果的な学習だと思いますか。	
	14. 他教科の学習において, 「思考した内容を言語化すること」がありますか。	6件法 (1: まったくない, ー6: とてもよくある)
	15. 他教科の学習において, 「思考した内容を言語化すること」は効果的な学習だと思いますか。	6件法 (1: まったく思わない, ー6: とてもよくそう思う)
まとめ	16. 後期に行った「ライティング課題＋振り返り活動」を通して, 英語ライティングの力が向上しましたか。	6件法 (1: まったく思わない, ー6: とてもよくそう思う)

注. 斎藤・鈴木(2022)を参考に筆者が作成

第36回 研究助成

B 実践部門・報告Ⅲ・英語能力向上をめざす教育実践

ディスカッション活動において高校生の 発話の量と質を高める取り組み

研究者:久山 慎也 広島県／広島県立安芸府中高等学校 教諭

《研究助言者:和泉 伸一》

概要

本研究は、高校2年生1クラスを対象に、英語コミュニケーションⅡの授業においてディスカッション活動に関する指導を2か月間実施した実践報告である。指導においては、ディスカッションを進めるための定型表現、理由を具体的に述べるためのPREP、会話を継続するためのディスカッション方略が取り上げられた。指導の前後で行った生徒のディスカッション活動における発話を分析したところ、英語使用の割合が増加し、理由をより詳しく述べようとする姿勢が生まれ、言いたい表現が出てこないときにお互いに助け合う様子が見られるようになった。また、指導後に書いた振り返りの分析からは、生徒は英語の使用割合が増えたことにより、達成感を持ち、ディスカッション活動に対する不安や苦手意識を軽減させたことも確認された。

1 はじめに

筆者の勤務校には県内で唯一の国際科が設置されており、学校の大きな特色となっている。国際科では、勤務校が教育目標として掲げる、持続可能な地域・社会づくりや世界の平和と発展に貢献できる人材の育成に向け、英語での発信力の強化を重視している。1年次にはアクティブスピーキング、2・3年次にはディベート&ディスカッションⅠ・Ⅱを学校設定科目として開講しており、これらの授業では、グループディスカッションやディベート、あるいはプレゼンテーションといった活動を中心に授業が組み立てられている。生徒は授業を通じて、海外の姉妹校生徒とのオンライン交流で自らの考えを発表する経験を多く積んでいる。学校外の関係者を招いて年に2回開催される学校関係者評価委員会でも、こうした取り組みは高く評価されている。その一方で、評価委員からは取り組みが国際科のみで完結しているため、普通科の授業でも同様の内容をいかに取り入れていくかが課題であると指摘されている。

本研究では、こうした校内的課題の解決に向け、筆者が昨年度の3学期に普通科2年生1クラスにおいて実施したグループディスカッション活動についての実践報告を行う。ディスカッション活動は英語コミュニケーションⅡの授業内で帯活動として行い、指導を通して生徒の英語による会話を継続する力を伸ばすことと、主張の理由をより具体的に述べさせることを目標とした。

2 先行研究のまとめ

2018年7月に告示された『高等学校学習指導要領解説 外国語編』では、やり取りや即興性を意識した言

語活動が授業の中で十分に行われていないことが課題として指摘されている。今回の改訂では、その克服に向け、話す技能において「話すこと[やり取り]」「話すこと[発表]」の2領域が設定された。本研究で取り上げるディスカッション活動は、このうちの[やり取り]の領域に深く関わる取り組みであり、上記の課題解決につながるものである。

英語コミュニケーションⅡでは、「話すこと[やり取り]」において2つの目標が設定されているが、このうち目標アでは「日常的な話題について、使用する語句や文、対話の展開などにおいて、一定の支援を活用すれば、多様な語句や文を用いて、情報や考え、気持ちなどを詳しく話して伝え合うやり取りを続けることができるようにする」(文部科学省, 2018, p.61)ことが目標とされている。ここで言う「やり取りを続ける」とは考えや気持ちを自然な会話の中で伝え合うことであり、そのためには非言語的な表現も含めた様々な定型表現と、会話の継続のために必要な表現を活用できる「会話を継続する力」が重要であるとされる(同書, p.62)。

このような会話の継続のために必要な表現として、本研究ではディスカッション方略を取り上げた。齋藤(2023)は英語ディスカッション指導を扱った先行研究より、ディスカッション方略を「ディスカッションをしている際に必要になる定型表現、役立つフレーズ」(p.122)であると定義している。また、それらを網羅的に記述したものとして Kehe and Kehe(1998)を挙げている。Kehe and Keheは13のディスカッション方略(using rejoinders/ asking follow-up questions/ seeking and giving clarification/ using comprehension checks/ answering with details/ soliciting more details from others/ interrupting others during a discussion/ recounting something they have heard/ volunteering an answer/ helping the leader of a discussion/ expressing an opinion/ referring to a source when giving an opinion/ leading a discussion themselves)とその練習のための活動を教材として紹介し、方略指導の具体案を提示している。また、やり取りを継続するための方略は、Dörnyei and Thurrell(1992)においても、13の会話方略(conversation strategies)として示されている。Dörnyei and Thurrellは方略使用の場面をディスカッションに限定してはいないが、Kehe and Keheが紹介したusing rejoinders/ asking follow-up questions/ seeking and giving clarification/ using comprehension checksと同様の方略を取り上げており、扱われている表現は指導者がディスカッション活動を組み立てる際に参考となる。

実際に方略指導の効果を検証した研究としては、Bejarano et al(1997)、Lam and Wong (2000)、Naughton(2006)、Lam(2009)が挙げられる。後の研究でも度々引用されるBejarano et al(1997)では、イスラエルの高校生34名を対象に、8週間にわたって週3回ある60分授業の30%をディスカッション方略指導に充てた結果が報告されている。取り扱われた方略は、ディスカッションにおいて何らかの言語的支援が必要となった際に用いられるModified-Interaction Strategies(MIS)と、ディスカッションを円滑に進めてまとまりをもたせるために用いられるSocial-Interaction Strategies(SIS)の2種類である。このうち、MISにはchecking for comprehension and clarification, appealing for assistance, giving assistance, repairingの4つの下位方略が、SISにはelaboration, facilitating flow of conversation, responding, seeking information or an opinion, paraphrasingの5つの下位方略があるとされた。事前・事後テストとしてディスカッション活動がなされており、そのturn総数及びMIS、SISの出現頻度を調べたところ、処置群及び対照群においてturn総数に変化は見られなかったが、処置群においてはディスカッション方略を用いないturn数が減少し、MIS及びSISの使用頻度が増加したことが確認された。

次に、Lam and Wong(2000)は、香港の高校生58名を対象としてclarifying oneself, seeking clarification, checking one's understanding of other peopleの3つの方略について指導した結果を報告している。Lam and Wongは、それぞれの方略について、事後テストのディスカッション活動において有効に使用された例と、使用されたものの上手く機能しなかった例の両方を挙げているが、方略使用が成功しなかった原因として、生徒の英語力の不足やグループ内の協力的な姿勢の欠如があることを指摘している。

続いて、スペインの大学生45名を対象に8回のディスカッション方略指導を行った Naughton (2006) では、follow-up questions, requesting and giving clarification, repair, requesting and giving help の4つの方略が取り上げられている。事前・事後テストとして録画されたディスカッション活動を分析したところ、turn総数では処置群及び対照群に変化は見られなかったが、処置群では4つの方略全てにおいて使用頻度が増加した。一方で、対照群ではrepairのみに増加が見られ、requesting and giving clarification の出現頻度は減少した。

最後に、香港の中学生40名を対象に、5か月間で1回80分の指導を8回行った Lam (2009) では、7つあるメタ認知方略の中でディスカッション方略が指導されている。データとしては、事前・事後テストで行われたディスカッション活動の発話、活動終了後に個々の生徒に対して行った面接の内容、全体に行った方略使用に関する質問紙調査の回答が分析された。その結果、ディスカッションを始める前に活動で何をすべきかを分析する problem identification と、やり取りの途中でメンバーに支援を求める asking for help の2つの方略で使用頻度が増加したことが確認された。

これらの先行研究では、事後テストにおいて処置群における方略使用が増加したことが明らかとなっており、指導の効果が表れていると結論づけられている。その一方、国内でディスカッション方略指導を扱った研究としては、大学生に4枚の絵を使ってグループで物語を作成させ、やり取りで必要となる表現を指導した加藤 (2016) や、高校生に設定したトピックについてディスカッション活動を行わせた児玉・亀谷 (2021) の実践報告があるものの、指導により方略使用がどう変化するかをデータに基づいて調べた研究は見当たらない。よって本研究では、日本人高校生を対象に方略指導を行うことで、生徒がディスカッション方略を使えるようになるのかをデータに基づいて調べることにした。取り扱った方略は、先行研究を参考に設定した次の5つである。

- ・やり取りにおいて相手の考えや気持ちを確認するための Clarification Request (CR)
- ・自分の気持ちや考えが正確に理解されたか確かめるための Comprehension Checks (CC)
- ・発言時に言語的支援を求めるための Requesting Help (RH)
- ・相手に言語的支援を提供するための Giving Help (GH)
- ・やり取りを活性化するための Follow-up Questions (FQ)

また、ディスカッションを円滑に進めるために Dörnyei and Thurrell (1992) が提示した相槌表現 (reactions) やつなぎ言葉 (fillers) も指導内容に含め、あわせて学習指導要領解説にある「様々な定型表現」としてディスカッションを開始したり、まとめたり、考えを述べたり、反論したりする際に必要となる表現についても指導した。なお、今回の実践研究は筆者の勤務校における授業で行うため、倫理上の配慮から対照群は設定しなかった。

さらに、本研究では上述した会話を継続する力に加えて、主張をする際に具体的に理由を述べさせることを指導のポイントとした。その理由としては、まず生徒にはディスカッション方略や定型表現を用いてやり取りを円滑に運べるよう指導することが重要であるが、その次の段階として、伝える中身を意識させる必要もあると考えたからである。学習指導要領解説では、やり取りの際には「順序立てて分かりやすく表現したり、具体例を挙げたりするなどして、理由や根拠などを具体的にかつ詳細に述べる必要がある」(文部科学省, 2018, p.62) とされている。一方で、ディスカッション活動における生徒の発話を観察していると、理由づけをする際の説明が不十分であり、具体例を出すなどすればより相手に伝わりやすくなると感じられる場面もあった。よって本研究では、主張をする際には PREP (Point-Reasons-Examples-Point) 構造を意識して、理由に具体性を持たせるよう指導した。

以上の点を踏まえ、本研究で設定した調査課題は以下の4点である。

調査課題1

定型表現を指導することで、ディスカッション活動における生徒の英語使用の割合は増加するか。

調査課題2

PREPについて指導することで、生徒はディスカッション活動で理由をより詳しく述べるようになるか。

調査課題3

ディスカッション方略の指導により、生徒の方略使用の頻度は増加するか。

調査課題4

ディスカッション指導に対して生徒はどのような感想を持つのか。

なお、調査課題4は、定型表現、ディスカッション方略、PREPについての指導全体が生徒にどのように受け止められたのかを調べる目的で設定した。

3 調査

3.1 対象クラス

対象生徒は筆者が昨年度、英語コミュニケーションⅡを担当した2学年普通科1クラス28名(男子13名、女子15名)の生徒である。当該学年では2クラスを発展1クラス、標準2クラスの3つに分け、習熟度別授業を展開している。調査対象となったクラスはこのうちの発展クラスである。2学期末の12月に実施をしたGTECアセスメント版スピーキング分野における生徒のCEFRレベルの内訳は、A2.2が8名(28.6%)、A2.1が18名(64.3%)、A1.3が2名(7.1%)であった。よって、実践を始めるにあたり、大半の生徒のスピーキング力は、慣れ親しんでいる話題や活動についての簡単に直接的な情報のやり取りができるとされる初級レベルであると確認された。なお、本実践を行うにあたり、調査結果は本研究以外の目的で使わないことと、個人が特定されるような情報を第三者に提供しないことが筆者により口頭で説明され、対象生徒より書面にて参加の同意を得た。

3.2 事前・事後テスト

事前・事後テストでは各グループ4名で7つのグループを設定し、7分間のグループディスカッションを行わせた。設定したトピックは、事前テストが「田舎と都会とでは暮らすのはどちらが良いか」、事後テストが「制服と私服とではどちらが良いか」である。グループの編成は出席番号で固定された座席に基づいて行い、性別や英語力についての配慮は特に行わなかった。また、設定されたグループは指導期間を通じて固定された。

ディスカッションに先立ってトピックがスライドで提示され、生徒はじゃんけんで司会者を決めた後に活動を始めた。ディスカッションの様子は各グループに配置したICレコーダーに記録されており、テスト終了後には筆者が回収し、文字起こしを行った。

また、一連のディスカッション指導を生徒がどう感じたかを調べるために、事後テスト終了後の次の授業で、生徒に活動についての振り返りをB6判サイズの用紙に7分間で書かせた。

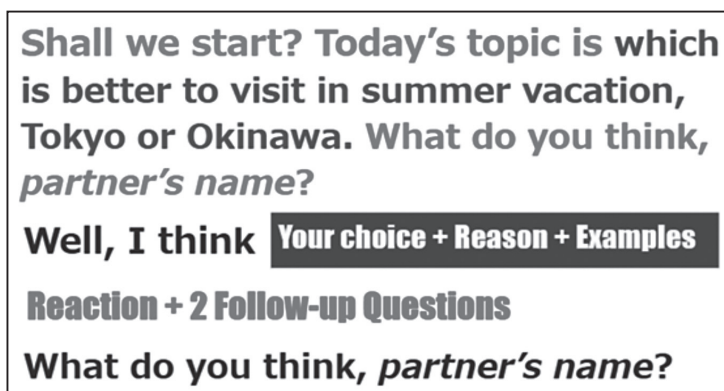
3.3 指導内容

英語コミュニケーションⅡでは50分授業を週4時間展開しており、今回の実践は2024年の1月から2月までの15回の授業で、冒頭の10分から15分を使用した帯活動として行った。各回の指導内容は表1に示したとおりである。

■表1: ディスカッション活動の指導内容

回数	活動内容	活動単位	時間	トピック
事前テスト	ディスカッション	グループ	7分	田舎と都会とでは暮らすのはどちらが良いか。
1	スピーチ	ペア	1分	ペットにするなら犬と猫のどちらが良いか。
2	スピーチ	ペア	1分	海外からの観光客に勧めるのは寿司とラーメンのどちらか。
3	会話	ペア	1分	お金と健康のどちらが大切か。
4	ディスカッション	グループ	7分	お金と健康のどちらが大切か。
5	会話	ペア	1分	ペットにするなら犬と猫のどちらが良いか。
6	会話	ペア	1分	生まれ変われるならどの動物になりたいか。
7	会話	ペア	2分	夏休みに遊びに行くなら東京と沖縄のどちらが良いか。
8	会話	ペア	3分	オンライン授業と教室での授業はどちらが良いか。
9	ディスカッション	グループ	7分	オンライン授業と教室での授業はどちらが良いか。
10	リスニング	個人	40分	モデルディスカッションの視聴(トピックは第9回と同一)。
11	会話	ペア	3分	田舎と都会とでは暮らすのはどちらが良いか。
12	会話	ペア	3分	タイムマシンがあれば過去と未来のどちらに行きたいか。
13	会話	ペア	3分	中学校では弁当と給食のどちらが良いか。
14	会話	ペア	3分	デジタル教科書と紙の教科書ではどちらが良いか。
15	会話	ペア	3分	春休みに行くなら東京と大阪のどちらが良いか。
事後テスト	ディスカッション	グループ	7分	制服と私服とではどちらが良いか。

基本的に生徒は提示されたトピックについて、自分の意見をペアやグループで話そう求められた。トピックは学習指導要領解説にある「日常的な話題」として、生徒にとって身近で学校生活や家庭生活と関わりがあり、立場を選んで意見を述べるようなものが設定された。第1, 2回の指導では生徒に話す量を増やすことを意識づけるため、西(2010)が考案したワードカウンターを使用し、提示されたトピックについてペアを変えながら1分間スピーチを3セット行った。第3回から第8回までは第4回を除き、生徒はペアでの会話を行っているが、会話を弾ませるためには良い聞き手の存在が必要であること、意見を述べる際には理由を具体的に述べるのが重要であると強調された。あわせて図1に示したようなスライドを用いて、生徒はReactions, Follow-up Questions, Comprehension Checksを意識的に用いるよう指導された。なお、第4回及び第9回の指導では4人のグループディスカッションを行い、議論を経験させた。



■図1: 授業で使ったスライドの一部

第10回の授業では、英語学習サイト www.engvid.com で公開されている、英語母語話者4名が「オンラインと教室での授業ではどちらが良いか」のトピックで行ったグループディスカッションの動画を視聴した。これは加藤(2016)の実践例を参考とした活動で、生徒は動画の視聴を通して自分たちのディスカッ

ションの進め方との違いや、使用されている表現を確認した。また、この回では資料1に掲載したフレーズ集も配布され、ディスカッションで利用できる表現や意見を述べる際のPREPの使い方も指導された。第11回から第15回の指導では、3分間のペアでの会話を、相手を変えて2セットずつ行ったが、活動の前にはフレーズ集で扱われている表現がスライドで提示され、生徒は毎回、口頭練習を行って定着を図った。

3.4 指導内容

3.4.1 発話データの分析

事前・事後テストで収集されたディスカッションの書き起こしデータについて、Stenström(1994)が提示した談話の階層構造(discourse hierarchy)の分析の枠組みに従い、まずはグループ毎に発話のturn数が数えられた。Stenströmはturnを“everything the current speaker says before the next speaker takes over”(Stenström, 1994, p.4)と定義しているが、相手からの発話が相槌(backchannels)である場合には、turnの移行が起こっておらず、前の話者のturnが継続しているとされる。また、会話の中で、生徒同士が笑い合う場面が頻繁に観察されたが、これはturnとしては扱わなかった。

次に、調査課題1について明らかにするために、各グループで、英語のみが使用されているturn数、日本語のみが使用されているturn数、両者が混在するturn数が数えられ、それぞれのturn総数に占める割合が算出された。あわせて英語のみのturnについては、平均語数も計算された。

続いて、調査課題2について調べるため、意見の理由を述べた箇所に当たるbecause節の平均語数が算出された。このbecause節の語数を数える際には、同じ語を繰り返してしまったり、言い直しをしたりする例が多く見られた。Ellis and Barkhuizen(2005)はこうした繰り返し(repetition)や言い直し(false start)は発話の流暢性を阻害する要因であると指摘しているが、本研究ではこれらの現象は、即興的な場面で生徒が発話を継続しようとするため生じていると肯定的に捉え、語数から除外はしなかった。

調査課題3については、事前・事後テストの書き起こしデータから5つの方略(CR, CC, RH, GH, FQ)に該当する箇所が特定され、その出現頻度が数えられた。なお、生徒の発話の中ではこれらの方略を母語で使用している例も見られたので、出現頻度は英語と母語それぞれに分けて数えられた。

3.4.2 自由記述データの分析

調査課題4について、ディスカッション指導を生徒がどのように感じたかを調べるため、生徒が自由記述で書いた振り返りが分析された。藤田(2019)は、生徒の自由記述回答を分析する際は、データを読み込みながら必要な箇所を抜き出し、気づきや考えたことをメモしながら当該箇所に内容を端的に表すコードを付与して、要約されたデータを整理するという手順を挙げている。本研究では、これらの手続きをより明示的・段階的に行えるSCAT(Steps for Coding and Theorization)を分析手法として用いた。SCATは大谷(2008, 2011, 2019)により開発された分析方法で、収集したテキストデータをセグメントに分割した後、それぞれに〈1〉データの中の注目すべき語句、〈2〉それを言いかえるためのテキスト外の語句、〈3〉それを説明するようなテキスト外の概念、〈4〉そこから浮かび上がるテーマ・構成概念の順にコードを考え出していくコーディング作業と、そのテーマ・構成概念を合わせてストーリーラインを記述した後に理論記述を行う作業からなる(大谷, 2019, p.271)。分析のためのマトリクスはExcelファイルとしてWeb上で公開されており、分析方法についても大谷(2019)において丁寧に解説してあるため、初学者にも着手しやすい。英語教育研究の分野でも今後利用の増加が見込まれる分析手法の1つである(三ツ木, 2020, p.54)。

今回、欠席者1名を除く27名より回収した生徒の振り返りは、総文字数3,876字であった。各生徒の振り返りに1つのセグメントを割り振った後、コーディング、ストーリーラインの作成と理論記述を行った。この分析の手順について具体的なイメージを伝えるため、グループ1のSCATで使用したマトリクスを表2に掲載する。なお、SCATでは手順〈5〉として、分析の際に感じた疑問や課題を書き込む欄が設けられているが、本研究では特に記載する内容がなかったため、マトリクスからは割愛した。

■表2: グループ1に対して行ったSCATのマトリクス

番号	発話者	テキスト	<1>テキスト中の 注目すべき語句	<2>テキスト中の 語句の言い換え	<3>左を説明 するようなテキスト 外の概念	<4>テーマ・構成 概念(前後や全体 の文脈を考慮して)
1	A	最初よりも話することができるようになった。他の人の意見に付け加えたり反対の意見を言うときにも英語で話すことができた。相槌やディスカッションを始めるときの言葉も覚えて使うことができた。	意見に付け加え/反対の意見/相槌/始めるときの言葉	追加/反論/バックチャネリング/定型表現	議論/ディスカッションにおける定型表現	定型表現の使用による議論の活性化
2	B	ディスカッションは苦手意識があったけれど、3学期に入り、何度もグループディスカッションを繰り返していくうちに少しずつではあるけれど、その苦手意識が和らいでいったと思う。	苦手意識/繰り返し/和らいで	不安/反復/軽減	話すことへの不安/成功への見通し	経験による苦手意識の軽減
3	D	最初はあまり言えなかったのとお題が難しくてどうしようとなっていたけれど、少しずつ言えることが増え、みんなに英語で何と言うか聞きながら意見を話すことができ、リアクションも少しずつできるようになった。同感を何と言うか忘れてしまったけど、みんながリアクションしたおかげでme tooだと思えることができた。	お題が難しく/言えることが増え/リアクション/みんな/思い出す	難易度高い/レパートリー増加/反応/メンバー/想起	タスクの難易度/Requesting Helpの習得/支援の提供	Requesting Helpの活用による相互支援の提供
4	C	最初は上手く喋れなくて黙ったりすることが多く、会話を弾ませることができなかったけど、だんだんと慣れてきて、お助けフレーズ集とかを活用していっぱい使ってみたり、相槌を頑張ったり、話を広げるなどいろいろできることが増えたので、ディスカッション練習はとても勉強になるし良い時間だったなと思った。わからない英単語があったら、すぐに友達を考えて答えてくれたりしたので流れがスムーズにできた。	黙ったり/慣れて/お助けフレーズ集/相槌/友達/スムーズ	沈黙/習熟/支援/バックチャネリング/メンバー/円滑	沈黙の克服/会話方略/相互支援	ディスカッション方略の活用による沈黙の克服

4

結果と考察

4.1 発話データの分析結果

表3は、グループ毎にディスカッションに要した時間、turn総数、英語turn数とその平均語数、日英混在turn数、日本語turn数、because節の平均語数をまとめたものである。

■表3: 各グループのturnの出現頻度及びbecause節の平均語数

		時間 (分秒)	turn総数 (回)	turn(英語)			turn(日英混在)		turn(日本語)		because 節語数 (語)
				数(回)	割合(%)	語数(語)	数(回)	割合(%)	数(回)	割合(%)	
G1	事前	7:00	24	4	16.7	8.3	5	20.8	15	62.5	8.3
	事後	5:37	28	14	50.0	8.5	10	38.7	4	14.3	16.4
	差		4	10	33.3	0.2	5	17.9	-11	-48.2	8.1
G2	事前	7:02	48	18	47.4	7.4	5	13.2	15	39.5	10.3
	事後	7:04	21	18	85.7	11.7	3	14.3	0	0.0	26.8
	差		-27	0	38.3	4.3	-2	1.1	-15	-39.5	16.5
G3	事前	6:55	77	28	36.4	3.8	20	26.0	29	37.7	10.0
	事後	7:05	59	40	67.8	6.3	16	27.1	3	5.1	33.5
	差		-18	12	31.4	2.5	-4	1.1	-26	-32.6	23.5
G4	事前	6:09	69	15	21.7	7.7	10	14.5	44	63.8	5.6
	事後	6:35	32	20	62.5	13.3	5	15.6	7	21.9	24.5
	差		-37	5	40.8	5.6	-5	1.1	-37	-41.9	18.9
G5	事前	6:14	61	36	59.0	3.5	15	24.6	9	14.8	12.8
	事後	6:57	29	23	79.3	8.9	2	6.9	4	13.8	18.0
	差		-32	-13	20.3	5.4	-13	-17.7	-5	-1.0	5.2
G6	事前	6:35	50	14	28.0	5.2	10	20.0	26	52.0	6.0
	事後	6:59	45	36	80.0	9.1	4	8.9	5	11.1	20.5
	差		-5	22	52.0	3.9	-6	-11.1	-21	-40.9	14.5
G7	事前	6:30	61	24	39.3	5.4	7	11.5	30	49.2	9.4
	事後	6:53	46	38	82.6	7.1	4	8.7	4	8.7	14.0
	差		-15	14	43.3	1.7	-3	-2.8	-26	-40.5	4.6

まず、調査課題1について、全てのグループのディスカッションにおいて、turn総数における英語turnの出現頻度の割合が事後テストで大幅に増加し、日本語turnの出現頻度の割合が大幅に減少していることが確認された。また、英語turnの平均語数についてはグループ間でばらつきがあるものの(0.2語から5.6語まで)、全てのグループで増加していることも確認された。

次に、調査課題2に関して、because節の平均語数はグループ間でばらつきはあるが(4.6語から23.5語まで)、全グループにおいて増加したことが確認された。これらの結果から、ディスカッション活動における定型表現の指導とPREPの指導により、生徒はやり取りにおける英語使用の割合を増加させ、理由をより詳しく述べるようになったことが分かった。

続いて調査課題3について、グループ毎に方略の出現頻度をまとめたものが表4である。全体的にディスカッション方略の出現頻度は多くはなかったが、グループ間での差が顕著であり、グループ3及びグループ6においては英語によるRHとGHの出現頻度が事後テストで増加したことが確認された。また、言語的支援を求めるRHについて、事前テストでは日本語による方略使用がほとんどのグループで見られたが、事後テストでは多くの場合、英語による方略に置き換わっていた。

■表4: 各グループにおけるディスカッション方略の出現頻度(回)

	開始時間／終了時間	Clarification Request (CR)		Comprehension Checks (CC)		Requesting Help (RH)		Giving Help (GH)		Follow-up Questions (FQ)	
		英語	日本語	英語	日本語	英語	日本語	英語	日本語	英語	日本語
G1	3:17 / 10:17 事前	0	0	0	0	0	1	1	0	0	0
	3:05 / 8:42 事後	0	0	0	0	0	1	1	0	0	0
G2	3:00 / 10:02 事前	0	0	0	0	0	3	2	0	0	0
	2:52 / 9:56 事後	0	0	0	0	0	0	1	0	0	0
G3	2:33 / 9:28 事前	0	1	0	1	0	2	1	0	0	0
	2:15 / 9:20 事後	3	0	0	0	4	0	7	0	0	0
G4	2:03 / 8:12 事前	0	0	0	0	0	1	1	0	0	0
	2:05 / 8:40 事後	2	1	1	0	2	0	1	0	0	0
G5	2:10 / 8:24 事前	1	0	0	0	0	1	2	0	0	0
	1:50 / 8:47 事後	0	0	1	0	0	0	3	0	0	0
G6	2:26 / 9:01 事前	0	0	0	0	0	0	2	0	0	0
	1:40 / 8:39 事後	0	0	0	0	6	0	7	0	0	0
G7	1:50 / 8:20 事前	0	0	0	0	0	2	0	0	0	0
	1:13 / 8:06 事後	0	0	1	0	2	0	4	0	0	0
計	事前	1	1	0	1	0	10	9	0	0	0
	事後	5	1	3	0	14	1	24	0	0	0
	差	4	0	3	-1	14	-9	15	0	0	0

4.2 自由記述データの分析結果

表5はSCATにより作成された各グループのストーリーラインであり、生徒が書いた振り返りの内容をグループ毎に圧縮したものである。なお、ストーリーライン中のかぎ括弧で囲まれた箇所は分析により付与されたコードを示している。

■表5: グループ毎に作成されたストーリーライン

G1	生徒は[ディスカッション方略の活用による沈黙の克服]と[ディスカッション方略の活用による相互支援の提供]及び[定型表現の使用による議論の活性化]を通じて[経験による苦手意識の軽減]に成功した。
G2	生徒は[協働学習による発話量の増加]や[協働学習によるスピーキング不安の軽減と有能感の獲得]とともに[定型表現の使用による議論の活性化]を経験することで[ディスカッション方略の習得による達成感]を得た。
G3	生徒は[Requesting Helpの活用による母語使用の減少]と[定型表現の使用による議論の活性化]や、[沈黙の減少による達成感の獲得]を経験した一方で[理由づけの難しさ]も感じていた。
G4	生徒は[多様なアイデア想起による足場掛け]による[理由づけの上達]と[Requesting Helpの活用による協働学習の進展と理由づけの上達]を経験し、[ディスカッションにおける流暢性の向上とまとめの難しさ]を感じていた。
G5	生徒は[噛み合わない議論の克服]や[語彙力の増加による意思伝達の成功]を通じて[議論のまとめ方の上達]を実感していた。
G6	生徒は[Requesting Helpの活用による不安軽減と議論の発展の実感]や[相互援助による母語使用の減少と英語使用の増加]を経験し、[定型表現の使用による沈黙の克服と英語使用への決意]を感じていた。
G7	生徒は[議論のまとめ方の上達]とともに[ディスカッション方略の習得による達成感と思考力の発達]や[定型表現の習得による流暢性の向上]を経験し、[定型表現の使用による議論の活性化とまとめの難しさ]も感じていた。

各ストーリーラインの記述より、生徒は定型表現やディスカッション方略を学習することでやり取りが活性化し、グループ内で協働的な雰囲気が醸成されたと感じていることが確認された。また、ディスカッションにおいて沈黙の時間が減少し、繰り返して練習することで活動に対する苦手意識や不安も軽減されたとの感想も持っていた。その一方で、一部の生徒は議論をまとめることに難しさを感じたことも分かった。

4.3 考察

調査課題1に関して、事後テストにおいて生徒の英語使用の割合が大幅に増加したが、その要因としては、振り返りにおいて多くの生徒が述べていたように、定型表現を練習したことで英語での議論の進め方に慣れたことが考えられる。Naughton(2006)も指摘するように、グループディスカッションでは即興的な発話が求められるため、議論を進めるために咄嗟に母語を使用してしまうこともある。実際、事前テストではそういった場面が多々観察された。英語による定型表現を身につけておくことは、こういった場面で過度の母語使用を回避するのに役立つ。また本研究では、こういった定型表現を指導するだけでなく、英語母語話者がディスカッションの中でこれらの表現を用いている動画を視聴させたことで、議論を進める上での見通しを生徒が持ちやすくなったことも考えられる。さらにグループのメンバーを固定したことで、お互いのことがよく分かり、協力的な雰囲気が生まれたことも、間違いを怖れず英語で話しやすくなった要因として考えられる。例えばグループ2の生徒は、振り返りの中で「回数を重ねるごとに話せる量が増えていったような気がする。相槌も最初より打てるようになって、リアクションがもらえると話していて楽しいと思うようになった。わからない単語や表現をグループ内で教え合いながら話すことができてグループ内の雰囲気も良かったから気軽に教え合えて良いなと思った」と述べており、安心して話し合う雰囲気があったことを報告している。ディスカッション活動の中で相互に助け合う雰囲気を作り出すことの重要性はLam and Wong(2000)やNaughton(2006)も指摘している点である。

次に、調査課題2について、事後テストではbecause節の語数がどのグループにおいても伸びていた。その要因としては、生徒にPREPを意識させたことで、理由を述べる際に具体的に説明する姿勢が出てきたことがある。また、理由を述べる際に、わからない単語があってもあきらめることなく、Requesting Helpを用いて発話を継続する意思を示せたことと、周りの生徒がGiving Helpや相槌を用いて発話者を

支援できたことも要因として挙げられる。ただし、繰り返しや言い直しによりbecause節が必要以上に冗長になってしまった面もあるので、つなぎ言葉を用いて上手に間をつなぐ技術も指導していく必要がある。

調査課題3に関して、今回の実践ではRequesting Help及びGiving Help以外の方略は出現頻度がそれほど増加しなかった。この点について、Naughton (2006) は同じ母語を持つ同年代の学習者同士の場合、言葉にしなくてもお互いに言いたいことが通じ合うため意味交渉が起こりにくく、Clarification RequestやComprehension Checksが使われにくいことを指摘している。しかし、対話相手が馴染みのあるクラスメートではなく、母語も年代も異なるような場合には、言わなくても分かり合えるということは期待できない。そういった際に、Clarification RequestやComprehension Checksを用いて意思疎通が上手くいっているか確認することは重要である。よって、この2つの方略については、事実確認を細かくしなければ完遂できないようなタスクを設定することで、その有効性に気づかせるような指導が今後必要となってくるだろう。最後に、Follow-up Questionsは方略指導の中でも度々取り上げたにもかかわらず、事前・事後テストにおいて全く観察されなかった。参加生徒がそれぞれの意見をどう伝えるかで手一杯となってしまう、述べた意見をさらに掘り下げていく段階にまで議論が至らなかったようである。今後、ディスカッション活動にさらに慣れていく中で、Follow-up Questionsの指導がより役立っていくのではないかと期待される。

調査課題4について、生徒の振り返りから、今回のディスカッション指導は生徒に好意的に受け止められたことが分かった。英語でのディスカッション活動に対する苦手意識や不安が軽減するとともに、英語が使えているという達成感やお互いに助け合えているという安心感を生徒は持ったようである。

5 まとめ

本研究では高校2年生28名を対象に2か月間で15回のディスカッション指導を行い、指導の前後でディスカッション活動におけるやり取りの変化とディスカッション方略の出現頻度の変化について調査を行った。事後テストの結果より、生徒は定型表現を活用することでディスカッションにおける英語の使用割合を増加させた。また、PREPを意識することにより理由づけの発話量も増やしていた。一方で、ディスカッション方略の出現頻度については、Requesting Help及びGiving Helpで使用回数の増加が見られたものの、Clarification Request, Comprehension Checks, Follow-up Questionsにおいては大きな変化は見られなかった。この3つの方略は、お互いに意思疎通が図れているか確認したり、議論を深めたりしていく上で必要となる方略であるため、その活用について生徒にさらに意識づけを行っていく必要がある。また、今回の実践では踏み込めなかったが、即興的な発話における正確性をどのように高めていくかは今後の大きな課題である。

指導後の生徒の振り返り文では、ディスカッション活動を継続することで成長や達成感を感じている記述も多く見られた。自らの主張について、言葉を尽くして相手に伝えられる生徒を育てるため、これからも取り組みを継続していきたい。

謝辞

本研究の実施に際して、鋭いご指摘と温かい励ましを与え続けてくださった上智大学外国語学部英語学科の和泉伸一教授に心よりお礼申し上げます。先生のご指摘により自分の研究を客観的に見つめることができ、いかに独りよがりな論を進めていたかに気づかされる場面が多々ありました。こうして直接ご指導をいただいた経験は、今後も実践研究を進めていく上で大きな財産となります。そして、このような貴重な学びの場を提供してくださった選考委員の先生方と公益財団法人 日本英語検定協会の関係者の皆様にも深くお礼申し上げます。最後に、授業に参加した生徒の皆さんにお礼を申し上げます。これからも良

い授業をすることが皆さんへの恩返しだと肝に銘じ、精進したいと思います。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

- Bejarano, Y., Levine, T., Olshtain, E., & Steiner, J. (1997). The skilled use of interaction strategies: Creating a framework for improved small-group communicative interaction in the language classroom. *System* 25 (2), 203-214.
- Dörnyei, Z., & Thurrell, S. (1992). *Conversation and dialogue in action*. Prentice Hall.
- Ellis, R. & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford University Press.
- 藤田卓郎. (2019). 「第6章 データを分析しよう」田中武夫・高木亜希子・藤田卓郎・滝沢雄一・酒井英樹 (編著)『英語教師のための「実践研究」ガイドブック』(pp.117-147) 大修館書店.
- 加藤和美. (2016). 「小グループが英語で打ち合わせ. 英語でプレゼンテーションでできる指導ーネイティブ・スピーカーのグループ活動から学ぶ」三浦孝・亘理陽一・山本孝次・柳田綾 (編著)『高校英語授業を知的にしたい: 内容理解・表面的会話中心の授業を超えて』(pp.198-228) 研究社.
- Kehe, D., & Kehe, P. D. (1998). *Discussion strategies: Beyond everyday conversation*. Brattleboro, VT: Pro Lingue.
- 児玉靖明・亀谷みゆき. (2021). 「『主体的・対話的で深い学び』による『思考力・判断力・表現力』の育成を目指した英語の授業づくりーグループ・ディスカッションを通してー」『朝日大学教育課程研究センター研究報告』23, 9-16.
- Lam, W. Y. K. (2009). Examining the effects of metacognitive strategy instruction on ESL group discussions: A synthesis of approaches. *Language Teaching Research*, 13(2), 129-150.
- Lam, W. Y. K., & Wong, J. (2000). The effects of strategy training on developing discussion skills in an ESL classroom. *ELT Journal*, 54(3), 245-254.
- 三ツ木真実. (2020). 「第3章 実証研究Ⅱ (質的研究)」廣森友人『英語教育論文執筆ガイドブックージャーナル掲載に向けたコツとヒント』(pp.43-58) 大修館書店.
- 文部科学省. (2018). 『高等学校学習指導要領解説外国語編・英語編』. 開隆堂出版.
- Naughton, D. (2006). Cooperative strategy training and oral interaction: Enhancing small group communication in the language classroom. *The Modern Language Journal*, 90(2), 169-184.
- 西巖弘. (2010). 『即興で話す英語力を鍛える! ワードカウンターを活用した驚異のスピーキング活動22』明治図書出版.
- 大谷尚. (2008). 「4ステップコーディングによる質的データ分析手法 SCAT の提案ー着手しやすく小規模データにも適用可能な理論化の手続きー」『名古屋大学大学院教育発達科学研究紀要 (教育科学)』54(2), 27-44.
- 大谷尚. (2011). 「SCAT: Steps for Coding and Theorizationー明示的手続きで着手しやすく小規模データに適用可能な質的データ分析手法ー」『感性工学』10(3), 155-160.
- 大谷尚. (2019). 『質的研究の考え方ー研究方法論から SCAT による分析までー』名古屋大学出版会.
- 齋藤英敏. (2023). 「英語グループディスカッション指導の理論と先行研究」『茨城大学教育学部紀要 (教科学)』72, 111-132.
- Stenström, A. B. (1994). *An introduction to spoken interaction*. Longman Group UK Limited.

Group Discussion お助けフレーズ集

ディスカッションを始める

Shall we start? Today's topic is ~. OK. What do you think about this topic, ~?

意見を言う I think ~. / In my opinion, ~

意見はPREP法で述べる!

Point → Reason → Example → Point

相槌を打つ I see. / OK. / That's true. / That's great. / Oh, yeah? / Really? /
That's too bad.

つなぎ言葉 Well / I mean / you know / kind of

発言しても良いか確認する May I say something?

発言を続ける意思を示す Oh, please let me finish.

分からない表現を尋ねる How can I say "~" in English?

メンバーの発言を確認する I'm sorry, but did you say ~? / You said ~, right?

メンバーにもう一度発言してもらう Could you say that again? / What do you mean?

メンバーの発言を拡げる Can I ask a question?

What ~? / Why ~? / How ~? / When ~? / Where ~?

メンバーの発言に付け足す You make a good point, and I would like to add that ~.

メンバーの発言に反対の意見を述べる That's true, but ~. / Yeah, but ~. /

You make a good point, but ~. / However, ~

ディスカッションをまとめる

So in conclusion ~. / Can we sum up and say ~?

第36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

用法基盤モデルに基づいたスピーキング指導が 中学生の『即興力』育成に及ぼす効果の検証

研究者: 吉澤 孝幸 秋田県／秋田県立秋田南高等学校 教諭・東京都／昭和女子大学大学院 在籍

《研究助言者: 小泉 利恵》

概要

本研究は、用法基盤モデルに基づいたスピーキング指導の有効性を検証することを目的とした実践研究である。用法基盤モデルの特徴である「語を核に文へ発展させる」という考え方は、メモによるキーワードから話すことと同じ視点をもつ。同時に、話す前のプランニングを効果的にするための指導を模索した。それは、令和5年に実施された全国学力調査[話すこと]（国立教育政策研究所，2023）における大問2の1分間の準備時間を生かすことにもつながるからである。

発話の前にキーワードだけによるメモを作成させることを実験群における「事前プランニング」と位置づけた。一方、日本において広く取り入れられている「リテリング」を比較群に取り入れた。比較群では、絵にキーワードを書き込むことを「事前プランニング」とし検証した。分析は、「事前・事後・遅延テストのルーブリック評価」，「スピーキングテスト前の1分間で作成したメモの質（評価）」，「これまで教科書で学習した固まり表現（軸語スキーマ）と自分のメモに書いたキーワードとを結びつけた数」を用いて分析した。

分析の結果、両群とも時間経過とともにルーブリック評価は向上した。両群の比較においては、実験群の方が、ルーブリック評価と「正確さ」「複雑さ」「流暢さ」の観点からの評価が有意に上回ったものの、その効果は遅延テストまでは持続しなかった。また、実験群の方がメモの質と発話の関係には強い相関が見られ、軸語スキーマとキーワードとの結びつけ数も有意に多かった。このことから、実験群によるメモの作成というプランニングを通して、発話のための認知資源の節約につながり、そのことが3観点（「正確さ」「複雑さ」「流暢さ」）の向上に寄与した可能性がうかがえた。同時に、メモの質が良くても、発話に結びつかないケースも見られ、その学習者の語彙の広さや深さにより、プランニングが効果的に働かないことも示唆され、今後の指導改善の指針を得た。

1 はじめに

教室においてスピーキングは、教育活動の基幹をなす技能である（Kurniatillah et al., 2022）。授業では教師が生徒に働きかけ、生徒がそれに反応し、続いて教師は生徒の反応にフィードバックを与える。このサイクルは、教室での教育活動を構成する最小の単位と言える。この単位が連続して授業が構成されることを考えると、スピーキングが基盤となり授業が運営されることは誰もが認識するところである。また、授業において読んだり聞いたりして受信した内容を、自分で再文脈化（田島，2001）し、テーマについての考えを他者と共有するためにも、スピーキングの技能は極めて重要である。

日本における英語教育政策においても、平成29（2017）年に告示された中学校学習指導要領（文部科学省，2017a）では、「話すこと」に関わる領域の目標の一つとして「社会的な話題に関して聞いたり、読んだ

りしたことについて、考えたことや感じたこと、その理由などを、簡単な語句や文を用いて述べ合う(話す)ことができるようにする」(話すこと[やり取り]・[発表], pp. 25-26)という目標が示されている。その目標を受けて『中学校学習指導要領(平成29年度告示)解説 外国語編』(文部科学省, 2017b)では、限られた時間に、まとまった内容を整理して伝えることや、自分で作成したメモを活用しながら話す活動を通して、即興で話したり、伝え合ったりする力をつけることを求めている。

しかしながら、スピーキングは複雑で難しい技能でもある(Burns, 2016)。その理由として、「話すこと」は「時間的制約のある処理(real-time processing)」(Kawauchi, 2005)に強く影響されることが挙げられる。限られた時間で、「概念化」・「言語形成」・「調音」・「自己モニタリング」という異なる作業を同時並行的に行わなければならない(Kormos, 2006)ところに難しさがある。特に中学生にとっては、限られた時間で同時並行的に複数のプロセスを処理することは、非常に難しい。令和5(2023)年に実施された全国学力調査「話すこと」における大問2は、やり取りを踏まえて「社会的な話題に対して自分の意見や理由を述べる」という内容であった。自分の考えを話す前に1分間で考えを整理する時間が与えられ、作成したメモを参照しながら話すことが許容された。大問2の正答率は4.2%であり(国立教育政策研究所, 2023)、その正答率からも目標とする生徒の現状との開きが読み取れる(国立大学法人 横浜国立大学, 2024)。このことから、学習指導要領が求める力と実際に教室で展開されている指導とを橋渡しする手段を導入する必要があると考えられる。これまで、スピーキングに対応する指導の視点としては、ロールプレイ、ドラマ、ディベート、ペアワークなどのタスク形式に多くの注意が払われていた(Hidayah, 2016)。しかし、産出のプロセスそのものに焦点が当てられた指導が十分行われてきたとは、必ずしも言えないのでないだろうか。

本研究では、スピーキングにおける「中学生の即興力育成」のための手段として、用法基盤モデル(Tomasello, 2003)の考え方を取り入れた。キーワードによるメモ(文部科学省, 2017b)で構想を膨らませた上で話すことは、用法基盤モデルにおける軸語スキーマに他の様々な語を結合させて文に発展させるという言語獲得についての考え方(橋本, 2017)と軌を一にすると判断したからである。用法基盤モデルに基づき、自分で作成したキーワードによるメモから話す指導を、本研究では「メモ式スピーキング」と呼ぶこととする。

本研究では、「メモ式スピーキング」を用いた指導の効果を検証するために、「メモ式スピーキング」を用いた実験群に加えて、比較群を設けた。比較群では、絵を用いたりテリングを取り入れた指導を行った。リテリングは、スピーキング能力の向上に効果的であり、コミュニケーション能力を伸長させる機会となると指摘されている(Putri et al., 2021)。その理由として、英語を第二言語として学習する場合や、外国語として学習する場合においても、読んだ内容を自分の言葉で表現することは、豊富な言語形式を取り入れることができる優れた方法であるとされていることが挙げられる(Zuhriyah, 2017)。また、近年日本の英語授業においても広く取り入れられている方法でもある(佐々木, 2020)。

2 先行研究

本節は、まず「メモ式スピーキング」が基づく理論や先行研究を述べ、即興で話すための手段としてメモをどのような観点で作成するかについて述べる。

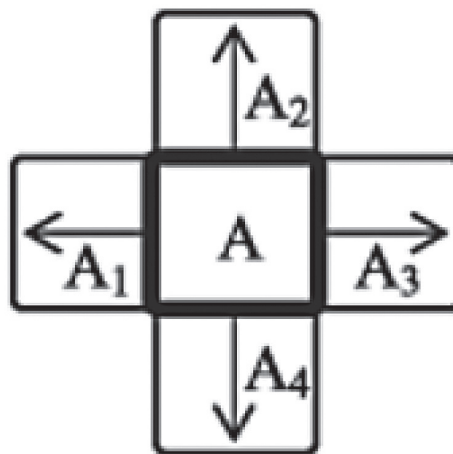
2.1 用法基盤モデルの言語習得観 —「メモ式スピーキング」との関係性—

本節では、用法基盤モデルの言語習得観と、「メモ式スピーキング」がどのように結びつくかについて先行研究を基に考察する。近年、人が持つ一般的な認知能力に依拠する用法基盤モデルからの言語習得が提唱されている。用法基盤モデルは、Langacker(1987)によって提唱され、Tomasello(2003)によって実証的に展開された言語習得理論である(富岡, 2016)。この用法基盤モデルは、生成文法のルールに依拠する

言語習得観と相対し、言語経験を通して事例からボトムアップで習得される(橋本, 2017)という点が大きな特徴である。用法基盤モデルにおいては、軸語スキーマと過去に使ったことのある表現の断片を「切り貼り(cut & paste)」し、発話における文を産出するとしている。軸語スキーマとは、軸とする語を中心とし、他の語と結びつく穴(slot)が備えられた固まりである(Tomasello, 2003)。すなわち、軸語スキーマとは、1語を指すのではなく、入れ替えが可能な部分が含まれた固まりを指すとしている(橋本, 2017)。そして、自分で過去に使用した固まりと固まりを切り貼りして文を産出するプロセスをusage-based syntactic operationsと呼んでいる(Tomasello, 2003)。また、用法基盤モデルにおける「切り貼り」には、二つの方法がある。一つ目は、軸語スキーマに含まれるスロットに他の語を入れる「fill in」である。二つ目は、既に想起している固まりの前か後ろに別の固まりを付け足す「add on」とあるとされる。すなわち、何度も使われることで慣習化した固まりを言語データとして蓄積し、それらの断片的なデータを切り貼りすることで発話を行うというプロセスと説明している。

その「切り貼り」を行う認知的なプロセスでは、発話場面に応じてコミュニケーションの目的を達成できるように軸語スキーマを想起し、その一部であるスロットに過去の言語経験で学んだ語を入れて、発話が行われる。例えば、子供が牛乳をもっと欲しい場合にMore milk. という表現を自分で蓄積した言語データから想起する。また、自分の頭にすでにある知識だけでなく、対話の相手が直前に言ったことも含まれるとしている。More milk. という表現を何回も使用していく過程で、異なる場面でMore X.のように一部がスロットになる。そして、More juice. という発話につながる。ここで重要なことは、More X. という軸語スキーマのスロットであるXにjuiceだけを入れ替えて産出されるのではなく、何度も使用して定着した固まり(軸語スキーマ)であるMore milk. がそのまま軸として想起され、生かされ続けるという点である。つまり、More milk 自体が、発話の軸語スキーマとして想起されるという認識である。そして、言語体験を通して獲得した軸語スキーマに内容を表す断片を組み合わせて表現するのが実際の言語運用であるとしている。

図1では、Aが軸語スキーマ、A1～A4がAの軸語スキーマを使った具体的な言語体験をしている。事例となる言語体験の内容はその都度変化するが、そこで使用する軸となる固まりは同一であり、その繰り返しの経験がAを定着させる。定着の度合いが強まることが枠線の太さによって表現されている。



■図1: 軸語スキーマを獲得していく過程(Langacker, 2009, p. 630のFig.2 (b)より)

図1が示すとおり、用法基盤モデルでは、言語経験により習得が進むと考えられている。そして、インプットは学習のための重要な言語材料と見なされる(橋本, 2017)。これは、最初から自分で使おうとする表現の固まりを考えるだけでなく、周りの人とのやり取りを通して対話者と文脈を共有した上で固まりを

取り込むものとし、それは共同注意 (joint attention) と呼ばれている (Tomasello, 2003)。このことは、インタラクションを通して、言語を使用する場面を理解し、それに応じた機能を持つ軸語スキーマを呼び起こしてくるものの必要性を述べたものである。本実践では、この軸語スキーマがメモとして書かれたキーワードと結びつくことが期待される。橋本(2017)は、軸語スキーマのことを、軸となる語とそれと結びつく穴であるスロットから構成される固まりとして捉え、「スロット付きスキーマ」と表現している。Tomasello(2003)によると、軸語スキーマは機能カテゴリー (functional category) に位置し、発話場面で必要とされる機能に基づき想起されるものとし、スロットは、内容カテゴリー (substantiative category) として自分が伝えたいメッセージであるとしている。さらに、これらの両者を包括するものとして連想カテゴリー (paradigmatic category) があり、一つの語を起点にしてそれにまつわる語や句を連想してくることができるとした。つまり、書き出したキーワードを起点に、発話場面にふさわしい軸語スキーマを結びつけることで文を産出することができると理論づけられる。この過程は、場面に応じて機能と関わる軸語スキーマを呼び起こして、アイデアとしてのキーワードを埋め込むことであるとも言える。さらに、軸語スキーマを呼び起こしてくる時には単語を個別に呼び起こしてくるのではなく、プレハブ表現 (prefabricated expression) を呼び起こすとされる。プレハブ表現は、村端・村端(2020)の定義によれば、文法規則によって、その都度生成されるのではなく、固まりとして慣習的に使用される一連の語句や文とされる。つまり、メモに書いたキーワードと使用場面につながる機能をもつ固まりを結びつけることになる。この固まり表現のことを、Tomasello(2003)は、fixed concrete expression と表現している。また、発話の流れを談話構成の観点から skeletal template と定義づけており、メモが発話の骨格として skeletal template になる可能性を見ることができる。さらに、Tomasello(2003)は、chunk of language をコントロールできるようにすることで、結びつけるチャンクを fluent unit of speech まで高めることの重要性を説いている。以上のことから、発話者はやり取りを通して会話の文脈を認識し、文脈にふさわしい機能をもつ軸語スキーマを想起する。そして、その軸語スキーマに付属しているスロットにメモとして書いたキーワードを結びつけるという発話プロセスが成立する。同時に用法基盤モデルの特徴である使用頻度を上げることで、軸語スキーマとスロットを結びつける時間を短縮することができる。そのため、用法基盤に基づく「メモ式スピーキング」を通して、想起したキーワードの拡張が容易になることが期待できる。言語経験を基に心的辞書 (mental lexicon) に自由にコントロールできる固まりを追加していき、より限られた時間で状況に応じて「切り貼り」できるようになることで、ゼロから文を構成する必要がなくなる。そのことが話し手の注意資源を節約することにつながり (Vercellotti, 2017)、スピーキングにおける即興力の育成につながると考える。

2.2 プライミング (Priming)

用法基盤モデルにおける発話プロセスでは、軸語スキーマに付属しているスロットに語を入れる際、次の二つの場合があるとしている。一つ目は、自分の頭の中にすでにある表現の断片を入れる場合である。二つ目は、対話の相手が直前に言った表現を取り込む場合である。このように対話者が使用した表現や断片を話者も使用する点については、プライミング (priming) 効果に着目する必要がある。「プライミング (priming)」とは、心理学において先行して受けた刺激 (prime) に影響され、後続の刺激 (target) に対する反応が起こることと定義されている (小柳・峯, 2016)。プライミングは、日常生活の会話の中でも頻繁に観察されることである (小柳・峯) が、第一言語話者と同様に、第二言語話者も対話の中で前の話者が使用した統語構造を自分の発話にも取り入れる傾向があるとしている (McDonough & Chaikitmongkol, 2010)。例えば、対話者が受動態の文を使用した場合、第二言語話者は同じく受動態の文構造を使用する傾向にあるとされる (McDonough & Chaikitmongkol, 2010, p. 84.)。このように、言語習得においては、統語的なプライミングが根底にあると言える (Bock & Griffin, 2000)。また、文レベルのプライミングの影響は、プライミング実験の直後にも見られると同時に、数週間後のポストテストにおいても観察された

(McDonough & Chaikitmongkol, 2010) という報告がある。さらに、第二言語話者にとって、効果的なインプットが与えられれば、母語話者が使用している構造に近い文を産出することにもプライミングは働くと述べられている(森下, 2017)。以上の見解から、キーワードでメモを作成させるにあたり、用法基盤モデルで提唱されている軸語スキーマを教師と生徒とのインタラクションに取り入れ、暗示的に与えておくことによって、プライミング効果が期待される。そして、呼び起こされた軸語スキーマとメモにあるキーワードが結びつくことで、キーワードから文への拡張が容易になるとと思われる。

2.3 プランニング(Planning)と階層的なマッピング(Hierarchical mapping)

文部科学省は、「話すこと」の指導に際し、予め文章で原稿を書いてそれを暗記して発表するのではなく、限られた時間で自分の考えを整理して発表するという「即興力」の育成を求め、その手段としてメモの活用を促している(文部科学省, 2017b)。しかしながら、そのメモの作成に至る指導手順や、そのメモをどのように活用し、即興力育成に結びつけるかは述べられていない。たとえ学習指導要領の持つ性格が「大綱的基準」(文部科学省, 2017b)であるにしても、実際の教室で効果的に活用されない限り、英語教育改革の実質化にはつながらない。以上の観点から、話す活動の前にメモをすることの位置づけを明らかにする必要があると考えた。また、メモを活用して話すことによって、スピーキングのどのような側面に効果が期待できるのかを整理しておく必要があると言えるだろう。

日本の公立中学校に学ぶ中学生にとって、最初から全て即興でスピーキングを行うことは難しいこと(国立教育政策研究所, 2023)であり、橋渡しとなる手段が必要である。その手段として本研究では、話す内容の概念化をスムーズにするために、「キーワードによるメモ書き」を1分間で作成させ、それを参照しながら発話できることを学習指導要領下における「即興力」と規定した。それは、令和5(2023)年に実施された全国学力調査[話すこと]における大問2における条件と同じくすることにもなる。

話す前にメモを作成することは、プランニングに相当する(Ellis, 2005)。Ellisによれば、プランニングは、事前に行う Pre-task planning と活動中に行う While-planning に分類される。さらに、Pre-task planning は、同じ活動を事前に行う Rehearsal と、何をどのように伝えるかを練る Strategic planning に分類される。話す前にキーワードによるメモを作成することは、タスクの前に行われ、かつその後と同じ活動を繰り返さないことを踏まえると、Strategic planning に属する。Ellis(2005)によると、Pre-task planning の機会が与えられることにより、スピーキングにおけるパフォーマンスが向上するが、「正確さ」「複雑さ」「流暢さ」のどの側面に効果があるかについては、必ずしも一致を見ていない。Ellis(2005)は、Wendel(1997)やOrtega(1999)を引用し、学習者にどのような点に焦点を当てて Planning を行ったかをインタビューした結果、語法・文法的な事項よりも話す内容の大まかなイメージづくりについて Planning を行ったと報告している。一方、Strategic planning を用いることで、「流暢さ」が伸長したという複数の報告がある(Foster & Skehan, 1996; Skehan & Foster, 1997; Wendel, 1997; Ortega, 1999)。このことは、第一言語発話生成モデル(Levelt, 1989)、またそれを受けた第二言語習得における Kormos(2006)の発話生成モデルにおける「概念化(conceptualization)」の部分に、Strategic planning の効果が働くことが期待される(Ellis, 2005)としている。しかしながら、効果の程度は学習者のレディネスによることも指摘されており、プランニングの効果がスピーキングのどの側面に働くかという点は、使用したタスクの種類や対象とした学習者の特徴により必ずしも同一ではないと解釈される(Johnson & Abdi, 2022)。また、本研究の対象者が、日本の中学生という英語初級者であることを鑑みれば、日本の検定教科書に使用されている言語材料を活用した中で、どのようなプランニングが効果的か、そしてどのように活用すべきかを検証する必要があると考えた。

事前プランニングにおいて、メモの型を示し、それに基づいてプランニングを指示した研究に岩男(2001)がある。岩男(2001)は、事前プランニングとして、(1)題目に関してアイディアを階層的に配列した階層的な概念地図を作成させる群(概念地図群)、(2)題目に関して考えたことを箇条書きさせる群(箇条書き群)、(3)メモや下書きを禁じ、直ちに文章を書き始めるよう求める群(準備書きなし群)を設定し、文章を生成

させた。その結果、内容の理解し易さは、概念地図群の文章が最も高く、準備書きなし群と箇条書き群は同程度であったと報告し、事前プランニングの種類が文章の理解しやすさに影響したと述べた。このことから、作文における文章生成の場合、単に概念の関係を記述する図ではなく、情報の提示に階層性を持たせた図を用いることで、情報のグルーピングが容易になることが明らかになった。本研究で扱う技能はスピーキングであり、文章産出における形式の面で岩男(2001)とは異なる。しかし、情報を階層的に配列するメモ(hierarchical note-taking)を導入することにより、「主題(最も伝えたい点)」から詳細情報を補足するという英語特有の談話構成に沿う発話になることが可能になると考えた。階層的に情報が配列されたメモを作成することは、言いたいことを連想的、羅列的に書き連ねていく「知識語り(knowledge-telling)方略」を避け、主題を明確にして発話の流れを一点に向けて収束させることができると予測した。以上の指摘から、どのように事前プランニングを行い、授業においてプランニングに基づいた指導をどのようにするかで、実際の発話に大きな影響が出ると考えた。

3 本研究の目的

本研究の目的は、用法基盤モデルの言語習得観に基づいたスピーキング指導が、中学生の即興力の育成に与える影響について検証することである。本研究における「即興力」とは、限られた時間で自分の伝える内容を整理して、自分の考えや気持ちなどを「発表」することとする(文部科学省, 2017b)。また、発話においては、即興性と同時にスピーチの内容が聞き手に対して明晰に伝わるような談話構成で話す力も研究対象としている。すなわち、ただ単に1分間で考えをまとめて話すことが本研究で求める即興力ではなく、英語特有の論理構成である「結論→理由→具体例」のような論理的かつ直線的な談話構成を1分間で準備して発話することを即興力があることとする。

先行研究からは、事前プランニングを行うことで話す内容の大まかなイメージをつくる概念化や流暢さに効果的に働くことは分かっている。しかしながら、キーワードによるメモ書きをプランニングと位置づけた時、どのようなメモを作成すると効果的かは調べられていない。作文における階層的なメモの有効性は提案されているが、中学生を対象としたスピーキングにおける階層的なメモの有効性は検証されていない。本研究ではスピーキングにおいて階層的なメモが果たす役割を検証する。

日本人中学生を対象とし、用法基盤モデルに基づく「メモ式スピーキング指導」を受けた実験群の学習者の方が、統制群である「リテリングを通したスピーキング指導」を受けた群よりも、論理的で即興的なスピーキング力を育成できるかを比較・検証することを本研究の目的とする。また、どのような要素が両群の発話の特徴に影響しているのかを調べるために、メモの良し悪しを判定した「メモの評価」と発話の関係を検証の視点とした。また、メモとして書いたキーワードと軸語スキーマが結びついている部分の数と「正確さ」、「複雑さ」、「流暢さ」へ与える影響を検証した。本研究では、指導前後の3回のスピーキングテストに焦点を当て、以下の研究課題を設定した。

研究課題①

「階層的なメモに基づくスピーキング指導を受けた群」は、「リテリングによるスピーキング指導を受けた群」より、スピーキングテストの「ルーブリック評価」による論理性のスコアが上回るか？

研究課題②

スピーキングテスト前に1分間で作成する「メモの質(評価)」とスピーキングテストの「ルーブリック評価(論理性)」には正の相関があるか？

研究課題③

「階層的なメモに基づくスピーキング指導を受けた群」は、「リテリングによるスピーキング指導を受けた群」より、「正確さ」、「複雑さ」、「流暢さ」が高まるか？

『中学校学習指導要領(平成29年度告示)解説 外国語編』(文部科学省, 2017b)で述べられている「メモから話す」ことの実質化を図るとともに、リテリングという学習指導要領改訂前から広く取り入れられてきた手法と比べた上で、効果を検証することを意図し、本研究課題を設定した。

4 方法

4.1 対象者

本研究の参加者は、公立中学校2年生50名である。50名を実験群25名(メモ式スピーキング指導)と比較群25名(リテリングを通したスピーキング指導)に分け、検証授業を行った。検証授業の前に、与えられたテーマに対する意見を述べる事前テストを行った。実験群と比較群の平均値を比較するために、 t 検定を行ったところ、群間に有意差は認められなかった [$t(50) = 1.28, p = .20, d = 0.85$]。

4.2 研究のデザイン

それぞれの群における指導デザインを、表1に示す。実験群と比較群において、InteractionとPlanningの内容は異なる。実験群におけるInteractionでは、後でメモに書き出すキーワードと結びつけやすい軸語スキーマを暗示的に与えることを目的とした。さらに、Planningの方法は、教科書の内容を要約した上で、自分の意見を述べるためにキーワードを「メインポイント→補足1→補足2→補足3→まとめ」の流れで階層的に配列するよう指導した。一方、比較群におけるInteractionでは、内容をスムーズに再生させることや、再生した内容に意見をつけ足すようにできることを目的とした。Planningの方法は、教科書の挿絵や付属のピクチャーチャートにある絵を使用し、それにリテリングに必要な語を書き込ませた。各群とも、Interaction・Planningと組み合わせてメモ式スピーキングやリテリングを行わせた。メモ式スピーキングとリテリングともに、Interactionと組み合わせた上で、それぞれ一つの方法論として見なしている。なぜならば、他者とのインタラクションを十分に行うことで、発話で使える語がプライミング効果により想起され、メモに書いたキーワードを結びつけることができ、その結果、統語的操作が容易になると考えたからである。

■表1: 実験群と比較群に行った異なる指導内容とテスト

実験群	1. 事前テスト	5月
	2. Interaction+Planning→Keyword-Based Speaking (メモ式スピーキング)	6月
	3. 事後テスト	7月
	4. 遅延テスト	11月
比較群	1. 事前テスト	5月
	2. Interaction+Planning→Retelling (リテリング)	6月
	3. 事後テスト	7月
	4. 遅延テスト	11月

注1: Interaction+Planningは、実験群と比較群では異なる内容

注2: 指導的な介入は実験群、比較群ともに事前テストと事後テストの間のみに行った。

指導的な介入を行った後から事後テスト・遅延テストまではメモ式スピーキング指導、リテリング指導ともに行っていない。

4.3 教材

両群の指導にあたって、題材は中学校検定教科書を使用した。事前テストから遅延テストまでは3単元にわたって行われた。群間での異なる指導は、一つの単元のみで行った。異なる指導を行った題材は、人口知能(AI)の発達とそれに伴う将来の職業の変化を扱った単元である。二つ目の単元では、ホームステイを通して日本の習慣やコミュニケーションのスタイルの相違について取り上げられている。さらに、最後の単元では、ユニバーサルデザインを通して共生の社会を考える内容となっている。二つ目の単元と三つ目の単元では、両群にスピーキングに焦点を当てた特定の指導は行っていない。両群で異なる指導を行う際、スピーキング前のプランニングにおいては、実験群へメモを作成するためのA5サイズの白紙のメモ用紙(資料1)を、比較群へは教科書図書の付属教材である「ピクチャーチャート」にある挿絵が印刷されている用紙(資料2)をそれぞれ配付した。比較群では、中央に絵が印刷された用紙に、リテリングする際に必要であると判断する語を書き込むことを認めた。

4.4 指導の実際

4.4.1 単元計画

実際の指導にあたっては、両群において表2に記載する指導手順に従って行った。異なる指導を行った単元を取り上げ、指導の流れを記した。表1に示される指導過程の中で、実験群と比較群で異なる指導を行った(囲みの内容を参照)。これらの指導は、2群とも同じ教師(研究者)が行った。

また、両群とも単元終了時における到達目標は、「中学校学習指導要領解説」(文部科学省, 2017b, p.26)にある第2の1(4)「話すこと[発表]」ウ「社会的な話題に関して聞いたり読んだりしたことについて、考えたことや感じたこと、その理由などを、簡単な語句や文を用いて話すことができるようにする」に関連し、同解説(p. 65)において示されている「社会的な話題に関して聞いたり読んだりしたことから把握した内容に基づき、自分で作成したメモなどを活用しながら口頭で要約したり、自分の考えや気持ちを話したりする活動」ができることとした。

授業の目標設定においては、学習指導要領の目標を踏まえ「聞いたり読んだりした内容を要約した上で、自分の考えや気持ちを伝えること」を限られた時間で話すこととした。発話に際しては、項目の羅列になると聞き手にとって、話し手が最も伝えたいことは何かが明確にならない。そのため、最も伝えたいこと(メインポイント)を伝え、それを補足するために、理由や具体例を組み合わせる効果的に伝えることに留意させた。

■表2: 単元指導計画 囲み部分は、実験群と比較群において指導過程が異なる。

時	指導内容
1st	1. Oral introduction of “We use computers to do many things.” 2. Oral introduction of today’s story (Scene 1) 3. Checking the new words and reading aloud. [実験群：メモ式スピーキング] 4. Planning time : 階層的なメモを作成する。 教科書本文の内容の論理構成を考えながら「階層的なメモ」を作成する。 5. Keyword-Based Speaking What is this article about? という問いかけに対して、自分で作成した階層的なメモに基づいて、要点と自分の考えを伝える。

1st	[比較群：リテリング] 4. Planning time：絵を利用する。 教科書本文からキーワードを抜き出し、絵に書き込む。 5. Retelling What is this article about? という問いかけに対して、絵を活用しながら本文内容を再生する。 その際、自分の英語で言い換えることを推奨する。
2nd	1. Review activity: Interaction for Summary 2. Oral Introduction of “I am surprised to see this.” 3. Oral Introduction of Today’s Story (Scene 2) [実験群：メモ式スピーキング] 4. Planning time：階層的なメモを作成する。 教科書本文の内容の論理構成を考えながら「階層的なメモ」を作成する。 5. Keyword-Based Speaking What is this article about? という問いかけに対して、自分で作成した階層的なメモに基づいて、要点と自分の考えを伝える。 [比較群：リテリング] 4. Planning time：絵を利用する。 教科書の登場人物の対話の流れを示した絵に必要な語を書き込む。 5. Retelling What is this article about? という問いかけに対して、絵に基づいて本文内容を再生する。 その際、自分の英語で言い換えることを推奨する。
3rd	1. Warm-up and New expressions something hot → something big → something to eat など連想ゲームをウォームアップ活動とし、その流れから新出表現である、We have various things to translate. を帰納的に説明する。 2. Oral Introduction of Today’s Story (Read and Think ①－1) and checking the new words 登場人物が書いたレポートの概要を、インタラクションを通して捉える。 [実験群：メモ式スピーキング] 3. Planning time：階層的なメモを作成する。 インタラクションを通して発話の構成とそこで使える軸語スキーマを意識する。 (メインポイント → 補足 1 → 補足 2 → 補足 3 → まとめ) その上で、メモを作成する。 4. Keyword-Based Speaking 作成したメモを基にレポートの要点とそれを受けて自分の考えを伝え合う。 [比較群：リテリング] 3. Planning time：絵を利用する。 教科書の登場人物の対話の流れを示した絵に必要な語を書き込む。 4. Retelling 絵に基づいて本文内容を再生する。その際、自分の英語で言い換えることを推奨する。

4th	1. Interaction for Summary (Read and Think ①－2) [実験群：メモ式スピーキング] 2. Planning time：前時に作成したメモを活用する。 前時で提示した発話の論理構成や、補足と共起しやすい軸語スキーマを、インタラクションを通して呼び起こす。インタラクションを踏まえ、前時に作成したメモにあるキーワードと軸語スキーマを結びつけるなど発話計画を練る。 3. Keyword-Based Speaking キーワードでメモを作成し、レポートの要点とそれを受けて自分の考えを伝え合う。 [比較群：リテリング] 2. Planning time：前時で使用した絵を活用する。 教科書の登場人物の対話の流れを示した絵を基に発話計画を練る。 3. Retelling 絵に基づいて本文内容を再生する。その際、自分の英語で言い換えることを推奨する。
5th	1. Oral Introduction of Today's Story (Read and Think ②－1) [実験群：メモ式スピーキング] 2. Interaction for Note-taking インタラクションの中で、新たな論理構成を暗示的に提示する。 Certainly, A, but B. 3. Planning time：階層的なメモを作成する。 4. Keyword-Based Speaking 作成したメモを基にレポートの要点とそれを受けて自分の考えを伝え合う。 [比較群：リテリング] 2. Interaction for Planning Q-A を通して、内容の主要な点を引き出し共有する。 3. Planning time：絵を活用する。 教科書の登場人物の対話の流れを示した絵に必要な語を書き込む。 4. Retelling 絵に基づいて本文内容を再生する。その際、自分の英語で言い換えることを推奨する。 5. Reading aloud 自分で使いたい表現に焦点化して音読を行う。
6th	1. Introduction for Summary (Read and Think ②－2) 前時の概要を想起しながら、インタラクションを通してストーリーを確認する。 [実験群：メモ式スピーキング] 2. Interaction for Note-taking インタラクションの中で、新たな論理構成を暗示的に提示する。 Certainly, A, but B. 3. Planning time：前時に作成したメモを活用する。 前時に作成したメモで発話計画を練る。 4. Keyword-Based Speaking 作成したメモを基にレポートの要点とそれを受けて自分の考えを伝え合う。

6th	[比較群：リテリング] 2. Interaction for Planning Q-A を通して、内容の主要な点を引き出し共有する。 3. Planning time：前時の絵を活用する。 前時で使用した絵を基に発話計画を練る。 4. Retelling 絵に基づいて本文内容を再生する。その際、自分の英語で言い換えることを推奨する。
7th	・ Talking about the Unit AI と人間について自由に考えを述べ合う。

4.4.2 プランニング前のインタラクション

話す活動の前に行うプランニングの方法は、実験群と比較群で異なる。実験群では、メモ式スピーキングで使用するための階層的なメモを作成した。一方、比較群では教科書の内容を表した絵に語を書き込み活用した。しかし、用法基盤の考え方に基づくと、語から発話を拡張するという点において、事前のインタラクションが重要な役割を果たすことを踏まえる必要がある。したがって、実験群・比較群ともプランニングにつなげるためのインタラクションを行った。そして、プランニングとして、階層的なメモを作成することや、リテリングのために必要な語を絵に書き込む活動と、事前のインタラクションを組み合わせ、一つのプランニングと見なした。

プランニングを行う前のインタラクションの目的は、実験群・比較群それぞれ次のように設定した。実験群では、目標とする発話の流れや枠組みを暗示的に示すとともに、生徒とインタラクションをしながらアイデア(メモに書くキーワード)と軸語スキーマ(メモに書いたキーワードと共起する既習の固まり表現)を結びつけるプロセスに気づかせることを目的とした。他方、比較群では、読んだ内容の主要な点を引き出すために教師が質問を投げかけ、その問いに対する答えをつないでいくことでリテリングがスムーズになるように指導した。また、比較群では、教科書の表現を自分の言葉で言い換えるように指導した。

<実験群における指導の流れ>

①プランニングの前のインタラクション

Teacher: In the story, Asami went to a translation company on Career Day. She learned a lot from Ms. Tanaka. She taught Asami its upside—good points and downside—bad points. What comes from your mind when you hear the word “AI”?

Student: Convenient, useful, and helpful.

Teacher: Yes, actually, Ms. Tanaka says that AI translate sentences quickly. Thanks to AI, you might finish doing your homework quickly, right? However, Ms. Tanaka tells us that AI has the downside—bad point. It sometimes misses important things. What does Ms. Tanaka say in the textbook?

Student: Misunderstand meaning.

Teacher: Do you have such an experience?

Student: When I want the information about candy, AI gave me the information about rain.

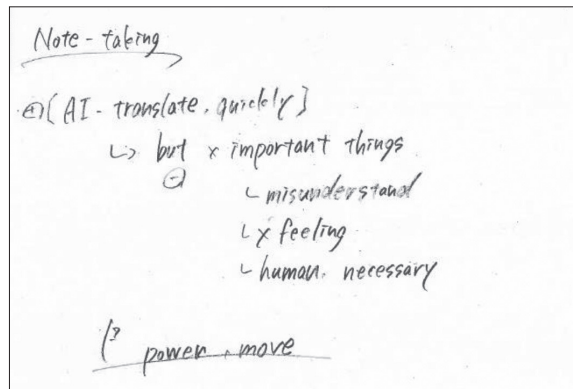
Teacher: Thank you for giving us a good example. So, we could say “Certainly, AI has good points. It can translate sentence quickly. But it misses important thing. Human is necessary. One example is misunderstanding the meaning. You could also say your experience at this point. Are there any downsides of AI?

Student: Not understand feeling.
Teacher: Exactly. Can you be happy when you communicate with AI?
Student: No.
Teacher: Can you be happy when you communicate with your classmates?
Student: Yes.
Teacher: I'm sure that your words can make your classmates happy by saying "Hello" or "Are you OK?"
They are short words, but we can be happy by them.
This means that human words have the power to move others.
The point here is that it is important to use AI effectively.
You could also say human words can change our feelings.
I cut "change our lives" in the first page of the textbook and then pasted it here.
To sum up, you could make a speech like this.
Certainly, AI is convenient, but it sometimes misses important things.
One example is to misunderstand the meaning. Another example is that it cannot understand our feelings.
This means that human words have the power to change our feelings.
The point here is that it is important to use AI effectively.

実験群におけるインタラクションの時間では、上記下線部のように生徒がメモで書き出すアイディア(キーワード)を埋め込むことができる軸語スキーマを暗示的に提示し、プライミング効果を期待した。また、その次の時間では、理由や利点を述べる時にどんな軸語スキーマを引き出せるかを問いかけて、意識化を図った(実際の授業では、もっと多様な生徒の体験などを取り入れたインタラクションを行った)。さらに、過去に学習した固まり表現(fixed concrete expression)を意図的に教師の発話に取り入れ、教科書の単元のどこにあった表現かという点についてもインタラクションを行った。その上で、生徒は教科書本文についてのメモを作成した。作成にあたっては、教科書に書かれてある内容を順番通りに追うのではなく、生徒自身が頭の中で情報を整理し、「メインポイント→補足1→補足2→補足3→まとめ」の流れを基本とし伝えることを求めた。

②メモを作成する(プランニング)

教師(著者)は、生徒とのインタラクションを進める際に黒板にキーワードを書きながら進めた。そのため、教師がティーチャー・トークに含めた軸語スキーマと自分のアイディアユニットであるキーワードを結びつけるイメージでメモを作成した。図2は、授業で実際に生徒が作成したメモである。なお、メモを作成する時は、題材である教科書本文を参照しながら作成させた。



■図2:メモ式スピーキングに使用する階層的なメモ(生徒による例)

③ メモに基づいて話す

図2のように作成したメモを基に、教科書の内容を端的に要約した上で自分の意見を述べる活動を行った。メモを基に話す活動は、一つのページにつき2時間を配当し、1時間目に5回、2時間目に7回実施した。また、2時間目の5回以降は、メモを参照せずに話す活動を行った。最終的には、メモがなくても即興で話せるようになることを目標にすることから、メモに基づいて話す活動の最後には必ず、メモなしで話す過程を取り入れた。

<比較群における指導の流れ>

① プランニングの前のインタラクション

Teacher: Now let's retell Asami's report. First, I'll give you some questions. But you don't need to answer the questions with full sentences. Just one word is OK. Do you think that AI is useful?

Student: Yes.

Teacher: When do you feel AI is useful?

Student: Shopping or studying.

Teacher: You mean that in shopping you don't have to use paper money or coins and you can use digital money, right?

Student: Yes. I use AI when I want some information.

Teacher: You can find the information quickly. Also, AI can change Japanese sentences into English sentences quickly. What did Ms. Tanaka say about AI?

Student: Sometimes miss important things.

Teacher: That's right. Repeat. "AI sometimes misses important things."

Student: AI sometimes misses important things.

Teacher: Good. Is AI good at understanding a writer's feelings?

Student: No.

Teacher: Do you think that humans need to check AI's performance?

Student: Yes. AI is not perfect.

Teacher: What did Ms. Tanaka say as an advice?

Student: It is important to use AI effectively.

Teacher: Why do we use language?

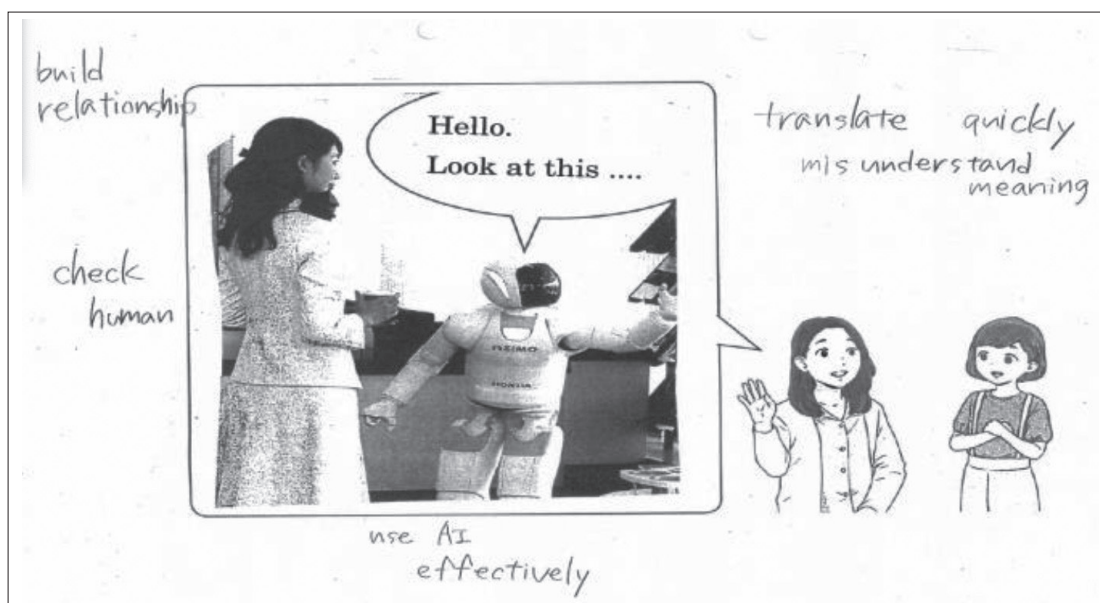
Student: To build relationships.

Teacher: Exactly. We can be happy when we build relationships by communicating with others. OK. Now let's retell the story with the picture chart. Please retell the story with these pictures on the board.

比較群では、上記のインタラクションに見られるように、英問英答を通して教科書の主要な点を引き出し、それらの語句をつないでいくようにした。教師は、引き出した語を板書しながら、教科書付属のピクチャーチャートに沿って書きながら全体の流れが視覚的につかめるように提示した。その後、黒板の絵や語から英文を組み立て、それを生徒にリピートさせ練習させた。なお、実験群で暗示的に与えられた軸語スキーマは、実践前の授業で実験群と同じ時期に、比較群には明示的に一度与えられた。具体的には、軸語スキーマで使用している表現(I think・・・や When SV, SV. If SV, SV. など)は、教科書の基本文であり、通常の授業中でも行われる文法知識の整理で扱われていたため、実験群だけが知っている表現ではない。しかし、比較群ではプランニングの前のインタラクションにおいて、その軸語スキーマを呼び起こす要素を教師の発話に入れなかった。

②絵に単語を記入する(プランニング)

発話の際に足場となる絵を基に、各自がリテリングを行う時に必要な語や、内容を伝えるために落としてはならないと判断した語を絵の周辺に書き込んで発話の構想を練った。下記の図3は、教科書の内容と関連する絵(教科書の挿絵や付属教材の絵・写真等)に、教科書本文から語を抜き出し、記入した例である。リテリングにおけるプランニングは、実験群と同じく、比較群でも教科書本文を参照して語を記入させた。



■図3:リテリングのための絵(授業で作成)

③絵を活用してリテリングを行う

語を記入した絵を活用して教科書の内容をリテリングする活動を行った。回数は、実験群と合わせ1時間目に5回、2時間目に7回行った。2時間目に行った際には、5回目以降は絵を見ないで内容を再生する活動を取り入れた。リテリングにおいては、書かれてある事実に加えて推論的な描写も積極的に行うように奨励した。

4.5 スピーキングテスト

両群とも指導の効果を検証するため、事前テストを行った。その後、6月に実験群と比較群に異なる指導を行った。6月の指導後は、メモ式スピーキングとリテリングを2群に分けた指導は実施せず、事後テスト・

遅延テストを行った。質問形式は、全国学力・学習状況調査(国立教育政策研究所, 2023)大問2の形式を参照した。スピーキングテストの実施時期とテストでの質問トピックは、表3の通りである。いずれもその時期に教科書の題材として取り上げられたテーマをトピックとしており、教科書の英文から得られた情報や表現を活用して応答することを目的とした。

■表3: 「スピーキングテストの実施時期と質問文」及び「収集したデータ」

実施時期	スピーキングテスト	スピーキングテストでの質問文と収集したデータ
5月	プランニング1分	What do you think of traveling abroad?
	↓ 事前テスト	収集データ ・ 生徒の発話を録音した音声データ ・ 音声データを手動で書き起こしたもの(トランスクリプト) ・ 生徒が1分間のプランニングで作成したメモ
7月	プランニング1分	What do you think of AI?
	↓ 事後テスト	収集データ ・ 生徒の発話を録音した音声データ ・ 音声データを手動で書き起こしたもの(トランスクリプト) ・ 生徒が1分間のプランニングで作成したメモ
11月	プランニング1分	What do you think of <i>Universal Design</i> ?
	↓ 遅延テスト	収集データ ・ 生徒の発話を録音した音声データ ・ 音声データを手動で書き起こしたもの(トランスクリプト) ・ 生徒が1分間のプランニングで作成したメモ

スピーキングテストは、日本人教師と生徒が対面で、1対1で実施した。スピーキングテストが始まる前に、テスト実施に関わる留意点と手順を次の指示文のように提示した。

- ・ 教科書のユニットで学習してきたテーマについて、これから英語で質問します。
- ・ そのテーマについて、どのように受け止めているか、あなたの意見や考えを述べてください。
- ・ その際、主張が明確に伝わるように、複数の補足を効果的に組み合わせてください。
- ・ 話す時間は、おおよそ1分30秒とします。時間になっても話している場合は、止めませんので最後まで話してください。
- ・ また、メモを作成して、必要に応じてメモを参照しながら話しても構いません。メモのための準備時間は1分です。

スピーキングテスト全体の流れは、次の順番で実施した。①簡単な挨拶とスモールトークを15秒ほど実施する。②指示文を文字で見せるとともに、英語の質問を見せながら話しかけをする。③メモを作成しながら、1分間で発話の準備をする。④トピックに応じて話す。話す時間は1分30秒を目安としていたが、時間が経過しても生徒がまだ話そうとする様子が見られた時は、発話を止めなかった。⑤教師が簡単な感想を英語で述べる。なお、3回の生徒の発話は全てICレコーダーに録音し、後日発話データは全て手動で書き起こした。また、1分間で発話の準備をした際に作成したメモは全て回収し、メモの良し悪しを判定し類型化するために回収した。なお、発話時間と発語数の記述統計は、表4の通りである。

■表4: 事後テスト(7月)における発話時間と発語数(記述統計)

群	発話時間		発語数	
	M	SD	M	SD
実験群	46.36	18.65	45.92	12.56
比較群	70.36	35.79	43.72	27.35

4.6 スピーキングテストと作成したメモの評価基準

スピーキングテストの評価にあたっては、「高等学校外国語科におけるパフォーマンステスト参考資料(指導者用資料)」(文部科学省, 2022)及びGTECスピーキングテスト評価基準(ベネッセコーポレーション, 2023)を基にして作成した(表5)。これまでの指導から、中学生が発話する際にはどうしても項目の羅列になりがちであることや、butを多用し、スピーチにおける主張が一点に向けて収束していないという問題を感じていた。そのため、GTECの評価基準では、Goal Achievement として「主張や意見を述べることを条件1とし、「それを効果的に伝えるために必要な補足情報の関係性が示されているか」という観点を条件2にしている点に着目した。また、「高等学校外国語科におけるパフォーマンステスト参考資料(指導者用資料)」からは、伝えなければいけないメインポイントを述べた上で、その補足となる理由などを詳細に述べているかという点を基準に設定しており、先のGTECにおける基準である補足情報の提示という観点と重なる点が多いと判断したことから、評価基準を作成する際の基資料とした。

■表5: 発話の評価基準

評価	言語(知識・技能)	内容(思考・判断・表現)
4	- 文の単位で聞き手の理解を妨げるような文法・語法の誤りがほぼなく、的確にメッセージを伝えている。 - 論理的な話の組立てをしており、多様な語句を組み合わせて使用している。	- 伝えたいポイントが述べられており、その上でアイディア間の関係性を考えて、効果的に補足が述べられている。 - 主張が一点に向かって収束しており、メッセージが明確、かつ効果的に伝わる。
3	- 文の単位で理解を妨げるような文法・語法の誤りが少しあるが、聞き手の理解に支障をきたすことなくメッセージが伝わる。 - 論理的な話の組立てをしている。	- 伝えたいポイントが述べられており、その上でアイディアを順序立てて補足が述べられている。 - 主張が一点に向かって収束しており、メッセージが明確に伝わる。
2	- 文として話しているが、文法・語法の誤りが多く、文レベルでメッセージが的確に伝わらない。 - 話の組み立てに一貫性がない。	- 伝えたいポイントは述べられているが、それを説明する補足情報が羅列になっている。 - 話の組立てに論理性がないため、話のポイントが明確に伝わらない。
1	- 発話の多くが、語句の羅列である。また、文形式で発話した場合も、文法・語法の誤りが多く、内容を推測することが難しい。 - 長い沈黙が多く、発話として成立しない部分がある。	- 語句の羅列などで話していて、伝えたいポイントに対する補足がない。 - 聞き手が推測しながら聞いても意味が伝わらない。

また、生徒の発話評価とメモの評価の関係性を分析するために1分間の準備時間で作成したメモは、表6の観点で評価し分類した。メモの評価にあたっては、英語のつづりの正誤は不問とした。また、Certainlyという完全な単語の代わりに、自分だけに判別できるシンボルとしてCと記すことはClippingと見なし、Note-takingのStrategyの一つであるためCertainlyと書いたことと同等に評価した。

■表6: 1分間で作成したメモの評価基準

評価	評価の観点
4	・英語の単語や語句だけで書いており、メインポイントと複数のサポートが書かれている。 ・メインポイントとサポートの関係性が階層的に書かれている。 ・最後の結論の部分にキーワードが配置されている。
3	・英語の単語や語句だけで書いており、メインポイントと複数のサポートが書かれている。 ・メインポイントとサポートの関係性が階層的に書かれている。
2	・英語の単語や語句だけで書いているが、メインポイントにサポートが入っていない。または、単に情報の羅列になっており、話の流れや方向性が明確でない。
1	・英語で書いているが、多くは文の形式で書いてしまっている。 ・日本語で書いている。

4.7 分析方法

本研究で使用した質問紙は無記名とし、得られた内容は本研究以外には使用しないこと、また、個人が特定されるような情報が第三者に知られることがないように暗号化を施した上で保管した。

収集したデータは、録音した音声データ、音声データを手動で書き起こしたもの(トランスクリプト)、及び生徒が準備時間に作成したメモである。録音した音声データ(参加者50人×3回)の評価は、日本人英語教師と英語母語話者が、「発話の評価基準」(表5)に従い、別々に行った。判断に迷った時は随時、英語母語話者の判断を参考にした。さらに、事後テストにおける「メモの評価」、事後テストにおける「軸語スキーマの利用数」「正確さ」「複雑さ」「流暢さ」のカウントも同じ2人の評価者が行った。信頼性の指標としてCohenのカッパ係数を算出した。Cohenのカッパ係数は「事前テスト」「事後テスト」「遅延テスト」「メモの評価」「軸語スキーマ利用数」の順に、.93, .84, .84, .94, .87となった。

トランスクリプトの評価・分析は「正確さ(accuracy)」、「複雑さ(complexity)」、「流暢さ(fluency)」の3指標(Foster et al., 2000)を使って行った。この3指標について、Koizumi et al.(2022)は、Housen, Kuiken, and Vedder(2012)を引用して、これらの3指標は第二言語の主要な構成要素であると述べている。生徒の発話を分析するにあたってはAS-unitを使用した。T-unitの使用も考えられるが、Foster et al.(2000)は、AS-unit(Analysis of Speech unit)では、プランニングにおいての統語的な面がより反映されると述べている。AS-unitは発話データを分析するために考案され、従属節や関係代名詞節などを含む独立節(independent clause)を1AS-unitとすると同時に、その中に含まれる従属節または節に準ずる句(sub-clausal unitと呼び、不定詞句や分詞句などが含まれる)を1節としてカウントする方法である(岩井, 2014)。そのため、トランスクリプトの分析は、AS-unitに基づいて言語的な分析を加えた。表7に各観点の指標を記述する。

発話の「正確さ」については、全てのAS-unit中で文法的誤りのないAS-unitの割合を算出した。また、スピーキングテスト前の1分間で作成した「メモの評価」、事後テストの「ルーブリック評価」、「正確さ」、「複雑さ」、「流暢さ」とのそれぞれの相関を分析した(ピアソンの積率相関係数 r を使用)。さらに、メモに書いたキーワードに軸語スキーマを結びつけている数と流暢さとの相関も分析した。

次に、「複雑さ」については、全体のAS-unitの中に含まれる従属節等の複雑な節または句を含む割合を算出した(例: 中学2年生を対象とするため言語材料が限られることから、if, when, thatの節を含む複文が中心となる。This means that SV. [1 clause, 1AS-unit]/ If you do something, you can do something. [2 clause, 1AS-unit]/ I'm sure that SV. [2 clause, 1AS-unit])。音声データのルーブリック

ク評価を数値化したもの、1分あたりの産出語数、全てのAS-unitの中での正しいAs-unitの割合、AS-unitの中に含まれる従属節等の複雑な節または句の含まれる割合、メモの評価とルーブリック評価との関係、そして軸語スキーマとスロットの結びつき数を用いて、二要因二元配置の分散分析(被験者間 [実験群・比較群]×被験者内 [事前・事後・遅延テスト])を行った。

「流暢さ」の測定方法については、「話した時間」「総語数」を用いた。語数を数える際には、トランスクリプトから繰り返し、言い淀み、言い間違いなどを全て除外した。これらの発話総語数を発話時間(秒)で割り、話者の流暢さを1分あたりの総語数(words per minute: WPM)で判断した。ポーズについては、Lennon (1990)の研究が知られるが、本研究の参加者である中学生の場合、発話の成立に関わるポーズの有無そのものが重要であり(茅野, 2018)、短い言い淀みまで含めることはしなかった。

表7: 3指標の一覧

「正確さ」	正しいAS-unit ÷ 全てのAS-unit
「複雑さ」	従属節や句を含むAS-unit ÷ 全てのAS-unit
「流暢さ」	話した語数 ÷ 話した時間 (秒) × 60秒

5 結果

5.1 スピーキングテストにおける二元配置分析の結果

実験群と比較群にそれぞれ、メモ式スピーキングとリテリングを実施し、その効果を検証した。メモ式スピーキングとリテリングの効果を測定するために行った事前テスト(5月)、事後テスト(7月)、遅延テスト(11月)におけるスピーキングテストの評価の平均値と標準偏差を表8に示す。

表8: 二要因混合分散分析の結果 (スピーキングテスト)

	実験群 (n=25)		比較群 (n=25)		時期			群			時期×群		
	M	SD	M	SD	F値	p値	ηp ²	F値	p値	ηp ²	F値	p値	ηp ²
事前テスト:5月	1.84	0.55	2.08	0.28	45.93	< .001	0.49	20.82	< .001	0.3	27.2	< .001	0.36**
事後テスト:7月	3.24	0.83	2.16	0.47									
遅延テスト:11月	3.44	0.91	2.36	0.63									

** p < .01, * p < .05

表9: 単純主効果

テスト時期	平方和	df	平均平方	F	p
事前テスト:5月	0.72	1	0.72	3.76	0.06
事後テスト:7月	14.58	1	14.58	31.93***	< .001
遅延テスト:11月	14.58	1	14.58	23.39***	< .001

群	平方和	df	平均平方	F	p
実験群	38.00	2	19.00	48.86***	< .001
比較群	1.04	2	14.58	3.57*	0.035

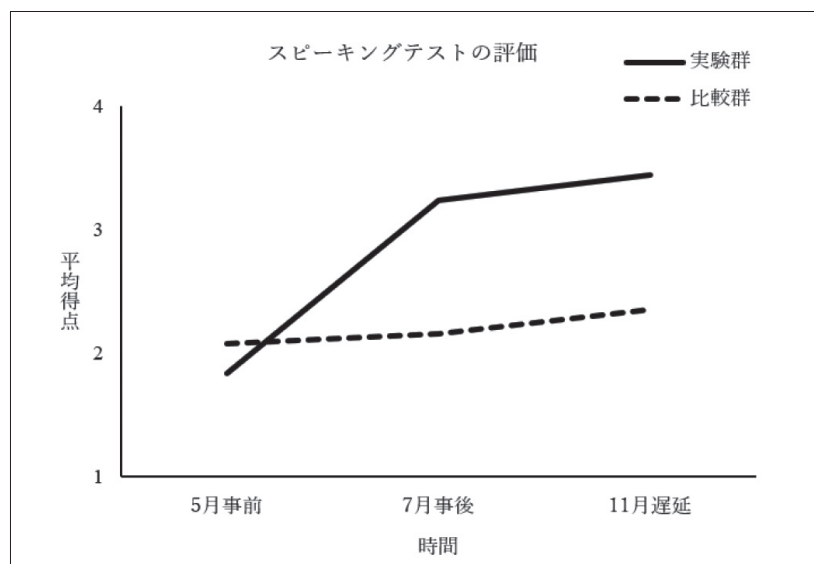
***p < .001 ** p < .01 *p < .05

指導法(実験群:メモ式スピーキング・比較群:リテリング)×スピーキングテスト実施時期(事前・事後・遅延)の二要因混合分散分析を行ったところ、指導法とテスト時期の主効果は、0.1%水準で有意で効果量もCohen(1988)に基づく大きい値であった(それぞれ、 $F(2, 96) = 45.93, p < .001, \eta p^2 = .49$; $F(1, 48) = 20.82, p < .001, \eta p^2 = .3$) (表8)。さらに、指導法とテスト時期の交互作用も1%水準で効果量も大きく有意であったため($F(2, 96) = 27.2, p < .001, \eta p^2 = .36$)、単純主効果の検定を行った(表9)。その結果、5月において群間でテスト結果の得点に有意な差は認められなかった($F(1, 2) = 3.76, p = .06$)。また、7月に実施された事後テストにおいて実験群と比較群の間で、ルーブリック評価の得点に有意差が見られた($F(1, 2) = 31.93, p < .001$)。つまり、7月の事後テストにおいてはメモ式スピーキングを取り入れた実験群の方が、リテリングを取り入れた比較群よりも得点が高かった。さらに、11月において実験群と比較群の間で、テスト結果の得点に有意差が見られ($F(1, 2) = 23.39, p < .001$)、実験群の方が、比較群よりも得点が有意に高かった。スピーキングテストにおける評価の推移を図4に示す。

■表10: 事後比較 群×スピーキングテスト時期

	95% 信頼区間				Cohen's d における 95% CI			
	平均値差	下限	上限	標準誤差	t	下限	上限	p
事前テスト:5月	0.24	-0.312	0.792	0.184	1.303	-0.487	1.224	0.697
事後テスト:7月	-1.08	-1.632	-0.528	0.184	-5.865	-2.647	-0.67	< .001**
遅延テスト:11月	-1.08	-1.632	-0.528	0.184	-5.865	-2.647	-0.67	< .001**

** $p < .01$, * $p < .05$



■図4: スピーキングテストの評価

5.2 「正確さ」「複雑さ」「流暢さ」二要因混合分散分析の結果

「正確さ」「複雑さ」「流暢さ」の3指標についても二要因混合分散分析を行ったところ(表11), 指導法とテスト時期の主効果は, 0.1%水準で有意であり, 中程度の効果量が見られた。([「正確さ」については, $F(2, 96) = 13.27, p < .001, \eta^2 = .22$; $F(1, 48) = 9.62, p < .001, \eta^2 = .17$;「複雑さ」については, $F(2, 96) = 14.83, p < .001, \eta^2 = .23$; $F(1, 48) = 5.69, p < .002, \eta^2 = .11$];「流暢さ」については, $F(2, 96) = 20.76, p < .001, \eta^2 = .3$; $F(1, 48) = 15.20, p < .001, \eta^2 = .24$)。さらに, 指導法とテスト時期の交互作用も1%水準で有意であったため, 単純主効果の検定を行った。その結果, 7月の事後テストにおいてのみ「正確さ」「複雑さ」「流暢さ」で実験群と比較群に有意な差が認められ, 実験群の方が有意に高かったことが分かった([「正確さ」 $F(1, 2) = 19.85, p < .01, \eta^2 = .30$;「複雑さ」 $F(1, 2) = 7.75, p < .01, \eta^2 = .14$;「流暢さ」 $F(1, 2) = 8.12, p < .01, \eta^2 = .15$)。しかし, 遅延テストまでその効果は持続しなかった。

■表11: 二要因混合分散分析の結果 (正確さ・複雑さ・流暢さ)

指標		実験群 (n=25)		比較群 (n=25)		時期			群			時期×群		
		M	SD	M	SD	F値	p値	η^2	F値	p値	η^2	F値	p値	η^2
正確さ	事前テスト:5月	0.43	0.25	0.53	0.26	13.27	< .001	0.22	9.62	< .001	0.17	19.85	< .001	0.3
	事後テスト:7月	0.78	0.21	0.41	0.27									
	遅延テスト:11月	0.79	0.21	0.57	0.24									
複雑さ	事前テスト:5月	0.17	0.27	0.19	0.16	14.83	< .001	0.23	5.69	< .002	0.11	7.75	< .001	0.14
	事後テスト:7月	0.46	0.21	0.22	0.17									
	遅延テスト:11月	0.37	0.23	0.31	0.15									
流暢さ	事前テスト:5月	35.95	15.10	34.75	14.31	20.76	< .001	0.3	15.20	< .001	0.24	8.12	< .001	0.15
	事後テスト:7月	63.59	16.04	41.21	19.29									
	遅延テスト:11月	53.41	17.20	38.33	13.20									

** $p < .01$

5.3 各指標の相関

実験群と比較群に異なる指導を行った後の事後テストにおける3指標([「正確さ」「複雑さ」「流暢さ」]及び「メモの評価」,「軸語スキーマとスロットが結びつけた数」)の相関を分析した(表12)。

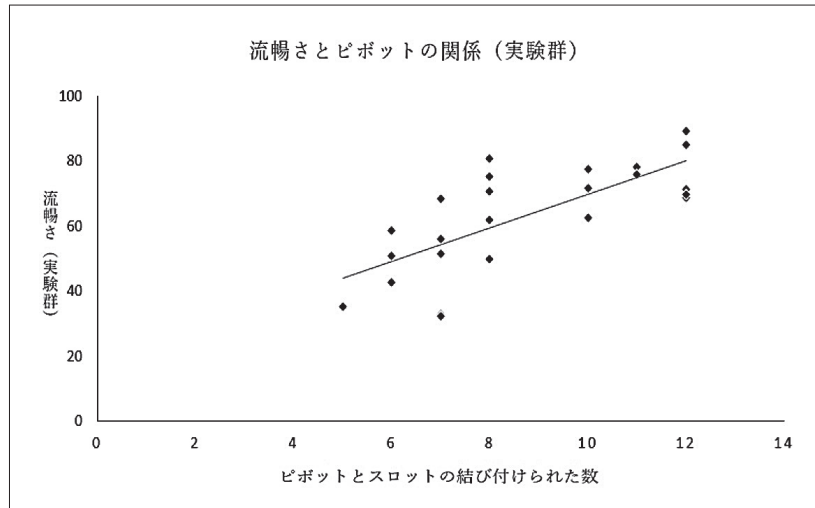
初めに, メモの質(4段階評価)と事後テスト(7月)の相関を分析した。分析の結果, 実験群ではメモの質(4段階類型)と事後テストの間には有意な正の相関が認められ, 効果量も大きかった($r = .78, p < .001, z = 1.05$)。また, 比較群においてもメモと事後テストの間には有意な正の相関は見られた($r = .44, p < .05, z = 0.51$)。実験群ではメモの質とスピーキングテストの評価に強い相関が見られた。1分間で行うプランニングと発話に強い関係性を見ることができた。

■表12: 事後テストに関わる要素の相関

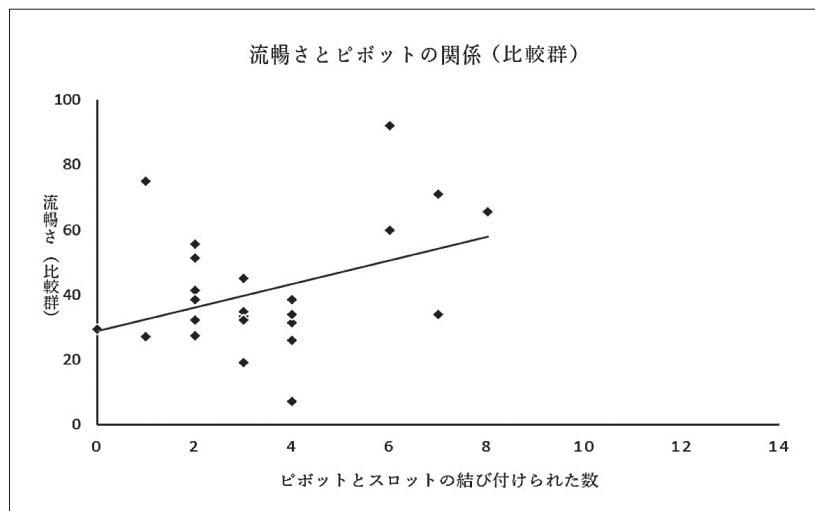
	1	2	3	4	5	6
1. メモ評価	-	.09	.44	.36	.08	.43
2. 結びつき	.20	-	.46	.58	.29	.37
3. 事後テスト	.78**	.06	-	.58	.52	.56
4. 正確さ	.53	.46	.67	-	.42	.45
5. 複雑さ	.27	.09	.29	.32	-	.09
6. 流暢さ	.20	.74**	.40	.30	-.03	-

** $p < .01$

さらに、軸語スキーマとメモに書いた語を結びつけている数と流暢さとの相関をプロット図に示す。実験群では、授業でのインタラクションを通して与えた固まり表現である軸語スキーマと自分のアイディアとしてメモに書いたキーワードを結びつけた数と「流暢さ」に高い相関が見られた($r = .74, p < .001, z = 0.94$)。一方、比較群では正の強い相関は見られなかった($r = .37, p = .006, z = 0.4$)。実験群と比較群におけるプロット図は図5・6の通りである。



■図5: 「軸語スキーマの利用数」と「流暢さ」の相関関係(実験群)



■図6: 「軸語スキーマの利用数」と「流暢さ」の相関関係(比較群)

6 考察

本研究では、用法基盤モデルの言語習得観に基づいたスピーキング指導を具体化したものとして、メモによるキーワードを拡張させて発話に結びつけるメモ式スピーキングの効果を、日本の英語教育において小学校から高校まで広く取り入れられているリテリング(米崎・松岡, 2022; 井長, 2019; 高木, 2022)と比較して検証した。初めに、設定した研究課題に沿って分析の結果を表13で整理する。

■表13: 研究課題と結果

研究課題	結果
研究課題① 「階層的なメモに基づくスピーキング指導を受けた群」は、「リテリングによるスピーキング指導を受けた群」より、スピーキングテストの「ルーブリック評価」による論理性のスコアが上がるか？	二元配置分散分析の結果、事後テスト（7月）と遅延テスト（11月）において実験群が比較群を上回り有意差が見られた。
研究課題② スピーキングテスト前に1分間で作成する「メモの質（評価）」とスピーキングテストの「ルーブリック評価（論理性）」には正の相関があるか？	実験群においては正の相関があり、相関係数も $r = .78$ と高い相関を示した。比較群においても正の相関は認められ、相関係数が $r = .47$ となった。
研究課題③ 「階層的なメモに基づくスピーキング指導を受けた群」は、「リテリングによるスピーキング指導を受けた群」より、「正確さ」、「複雑さ」、「流暢さ」が高まるか？	実験群と比較群を比較したところ、事後テスト（7月）のみ「正確さ」「複雑さ」「流暢さ」における群間差が有意に認められた。しかしながら、遅延テストまでその群間差は維持されなかった。

表13における結果を基に、各研究課題において考察を加える。最初に、メモ式スピーキングとリテリングにおけるルーブリック評価結果から考察する（研究課題①の検討）。5. における分析結果では、実験群と比較群とも経時的にそのスコアが伸びていることから、メモ式スピーキングとリテリングの両方がルーブリックの観点からスピーキング力向上に寄与した可能性がある。一方、事後テストと遅延テストにおいて、実験群のスコアの方が比較群よりも高かった。このことは、指導的介入を行ったことで、事後テストから遅延テストまでその効果が継続した可能性が示唆される。メモ式スピーキングにおいては、最も伝えたいことを左上に記入し、詳細情報になるに従い左下に配置するという階層的な約束事がある。この流れに沿ってアイディアを構成したり、話したりすることにより、何かを言ったら補足をいれるという発話の最小ユニットが実験群の生徒に根づいたと考えられる。発話生成モデルについてKormos (2006) では、(1) 概念化 (Conceptualization)、(2) 言語形成 (Formulation)、(3) 調音 (Articulation) (4) 自己モニタリング (Self-monitoring) の四つの構成要素からなると述べられている。そして熟達度の高い話し手は、これらのプロセスを同時並行的に処理することができるとしている。発話プロセスにあたっては(1) から(3) の順序で行われるが、Ellis (2005) は、タスクの事前に行うプランニングの効果は、概念化(Conceptualization)をしやすいことだと述べている。今回の実践においても、メモ式スピーキングにおけるメモの作成の仕方が、より論理的な発話の流れを誘発した可能性がうかがわれる。授業における指導を通して、メモを見ながら何回も使う言語体験を積みながら、最終的にメモなしでも頭に骨格が残ったと考えられる。

また、メインポイントを述べてから聞き手に納得してもらうために、異なる種類の補足を組み合わせて述べるようなキーワードを配置することも実験群では指導されている。そのため、「利点」を述べる時にはどのような軸語スキーマを呼び起こしてくれればいいのかを無意識的に考えたと分析できる。このことは、軸語スキーマを呼び起こし、自分で浮かんだキーワードと軸語スキーマを共起させた可能性を示唆している。遅延テストにおいても、実験群と比較群との間に有意な差が開いたことに着目したい。Tomasello (2003) では、スピーチの骨格(skeletal template)が、言語経験を通して形成されることの重要性が述べられている。一つのトピックを通して、発話の「骨格」を作ることで、別のトピックになっても活用できるようになることを指している。この骨格に沿った流れを意識して話すことで、発話内容が箇条書き的になることを避け、英語話者の論理に沿った発話に近づけることができると思われる。他方、比較群で導入したリテリングについては、絵があるため学習者にとって話しやすいことが利点である。流れをあまり気にせず、言えるところから「描写」することができるため、一定の発話数が確保される様子が授業からも観察された。表現形式についても、教科書の本文をそのまま再生するのではなく、自分の既習の語や表現で言い換えをすることを奨励してきた。この過去に学習した表現の固まりを軸語スキーマと見なした時、リテリングで

は「理由」「利点」「価値づけ」などの観点が一定の順序で出てきにくいこともあり、それらの観点と軸語スキーマを結びつけにくい面もあるのではないかと考えられる。Ortega(2005)は、熟達度の低い学習者はプランニングの際に単語や表現を引き出すのにフラストレーションを感じると指摘しており、学年も考慮した発達段階に応じて、描写の場合はリテリングを用いたり、意見表明の場合はメモ式スピーキングを取り入れたりするなど、両者を組み合わせていくことが効果的であろう。

次に、二つ目の研究課題であるメモの質(評価)とスピーキングテストの評価の関係について考察する。分析の結果にあるように、実験群と比較群をまとめた全ての生徒において、正の相関が見られた。中学生の場合、スピーキングテストでは一定の時間に、話す論点を一点に収束してまとめることは難しい。話している途中で思いついたことを加えたり、そのことで話の流れが拡散したりしてしまう場合も少なくない。情報が階層的に配列されたメモの作成を授業で指導されることによって、スピーキングテストの準備時間である1分間でも、同じような手法で効果的にプランニングができたことを示している。特に、メモの最後に、スピーチのまとめの印となる語を入れることで、自分のスピーチを論理的にまとめることにつながっていると判断できる。Ortega(2005)は、事前プランニングでは発話において伝えようとする主要な点などの内容面に注意を向けることを優先するものの、同時に言語形式にも注意を向けている。言語形式については、それまでに身に付けた心的辞書 (mental lexicon) から引き出すわけであるが、Ellis(2005)は心的辞書の中にある語彙的固まり (lexicalized chunks) が不足している場合、規則に基づいた文法 (rule-based grammar) に頼るとしている。このことから、授業における言語活動で十分な語彙的固まりの取り入れが行われた場合は、スピーキングテストにおいても作成したメモを活用できるといえる。しかし、語彙的固まりが不十分である場合は、メモとの共起がスムーズに行われず、ルールに基づいた発話に頼り、誤用や沈黙が生じてしまうことが考えられる。メモが指導された通りに作成できても、発話やそのルーブリックの評価スコアや流暢さなどに寄与しないケースもある。その場合は、メモと語彙的固まりである軸語スキーマとの共起が不十分であることが原因である可能性が高い。Koizumi and In'nami(2013)は、より広く深い語彙知識を備えている学習者は、使用しようとする語へのアクセスが容易になり認知資源を節約することにつながるとし、その分「概念化」や「言語形成」に注意を向けることができると述べている。これらのことから、語彙において広さや深さ、特に語彙的固まりの知識と使用する力を備えることが、メモによる「概念化」や「言語形成」もスムーズにし、プランニングをより効果的にする重要な要因になると考察できる。つまり、プランニングの方法だけに注力するだけでは十分とは言えず、メモに書いたキーワードと結びつけることができる語を増やすことも大切である。

三つ目の研究課題は、「正確さ」「複雑さ」「流暢さ」における群間の違いについてである。事後テストにおいては、「正確さ」「複雑さ」「流暢さ」とも実験群の方が比較群の数値よりも有意に上回った。このことから、正確さにおいては、メモ式スピーキングの指導を受けることを通して、軸語スキーマをより多く自分でメモに書いたキーワード(スロット)と結びつけたことが効果的に働いたと分析した。Tomasello(2003)が述べているプレハブ的表現 (prefabricated expression) で発話を組み立てたことにより、初めから文法ルールに基づいて発話を組み立てるより、ミスが少なくなったと思われる。「複雑さ」については、実際の発話のトランスクリプトを調べると、I think that S V. が発話の最初に使われたため、両群において複文の使用例の数が確保されていたが、それ以外の複文の使用は This means that S V., または、If S + V, S + V. など、軸語スキーマとして暗示的に教師がインタラクションの中で与えていたものがほとんどであったことが判明した。この軸語スキーマの使用が、次の観点である「流暢さ」にも影響したと言える。つまり、メモのキーワードに語彙的固まりである軸語スキーマを結びつけて発話をつくることで、より短い時間で発話が形成された可能性がうかがえる。このことは、軸語スキーマとキーワードとの結びつけ数の相関関係からも支持される。

事前プランニングが「正確さ」「複雑さ」「流暢さ」のどの側面に効果があるかについて、Ellis(2005)は「正確さ」よりも「流暢さ」や「複雑さ」により効果があるとしている。それは、どのように話すかという言語的な側面よりも、話そうとする内容に関わる青写真を描くことにより注意を向けるからだとしている。ま

た、事前プランニングは、「流暢さ」において最も貢献すると指摘されている(Foster & Skehan, 1996)。Johnson and Abdi (2022) は、事前プランニングが第二言語の発話に及ぼす影響の中で、最も強いものが「統語的複雑さ」で、次に「流暢さ」であるとしている。しかしながら、本研究においては指導的介入を行った後の事後テストでは、「正確さ」「複雑さ」「流暢さ」の全てにおいて比較群よりも実験群が有意に上回った。これは、本研究における対象となった言語材料がかなり限定されていることから、事前プランニングの効果が及びやすいことが理由として存在すると考察している。具体的には、使用する言語材料が中学2年生までの履修内容であることや、教科書におけるターゲットとした単元が一つであることなどが理由として考えられる。また、同じ事前プランニングを行っても、先行研究と同じ結果が出ないことについては、多くの先行研究ではプランニングの時間が10分または5分であるのに対し、本研究ではテストという場面もあり1分のプランニング時間であるという違いも影響を与えているからではないかと考えている。

7 まとめ

本研究を通して、用法基盤モデルの特徴である「語は構文である(word is structure)」という考え方が、日本の中学生を対象としたスピーキング指導においても有効に働く可能性を見ることができた。形態としては語であっても、それを構文としてみなすという考え方は、軸語スキーマと他の語を共起させることにより実現される。たとえ一語であっても、それを教室内で拾い、文に拡張していく。用法基盤モデルの大きな柱として“filling slots in pivot schemas”(Tomasello, 2003, p. 327)がある。場面や機能に応じた軸語スキーマを想起し、軸語スキーマに含まれるスロットに多様な語を入れる言語体験を通して、軸語スキーマが学習者の発話における「骨格」となり得る。この骨格は、より深く単語の知識を習得することでより結びつきが強くなり、発話における「正確さ」「複雑さ」「流暢さ」の向上につながると考える。そうすれば、定型句の性質を帯びる軸語スキーマにとどまらず、様々な語が軸語スキーマとなり、他の語との多様な共起ができるようになることが期待される。

メモ式スピーキングで形成された発話の骨格(skeletal template)を紙に書かずに、文脈に応じて呼び起こしてることができることが、即興力の育成につながっていく。そして「学習指導要領解説」(文部科学省, 2017,b)に示されている「メモ」の指導方法と効果について、一つの提案としてより多くの英語教師と共有できることを期待したい。一方、今回の結果はタスク形式が「与えられたテーマについて意見を述べる」という固定されたものであることから、今回使用したテーマに限定される可能性もある。そのため、ここからどの程度まで一般化できるかは今後の課題である。

本研究では、Ortega (2005)の言う「学習者の受け止め(Learners' perception about planning)」を通しての分析は行われなかった。今後、メモ式スピーキングが、発話準備または発話におけるどの側面に効果的であるかについて学習者の認識を調査する必要がある。同時に、今後学習者の持つレディネスとの関係も分析していきたい。用法基盤モデルに基づく指導においては、特定の文構造を使用する「頻度」が重要な要素であることから、今後メモを基に話す言語経験を行う頻度についても研究を進めていく必要がある。そして、最終的にはメモを参照しないで話す「即興力」を育成するために、メモという「足場」をどのように外していくかも今後の課題である。

謝辞

本研究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様と選考委員の先生方、特に研究助言者をお引き受けくださった小泉利恵先生には、研究デザインの段階から報告書の執筆までにわたり、丁寧なご指導をいただきましたこと、心より感謝申し上げます。

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

- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of experimental psychology: General*, 129(2), 177-192. <https://doi.org/10.1037/0096-3445.129.2.177>
- Burns, A. (2016). Research and the teaching of speaking in the second language classroom. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 3; pp. 242-256). Routledge.
- Ellis, R. (2005). *Planning and task performance in a second language*. John Benjamins.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3), 299-323.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354-375.
- 橋本ゆかり. (2017). 認知言語学・用法基盤モデルの誕生の背景からみる特徴と第二言語習得研究への応用. 横浜国大言語研究, 35, 151-163. <https://ynu.repo.nii.ac.jp/records/8094> ベネッセコーポレーション. (2023). GTEC Speaking テスト・Writing テスト 採点基準 (Vol. 3). <https://www.benesse.co.jp/gtec/fs/scoring/>
- Hidayah, R. S. (2016). *Improving students' speaking skill through retelling technique using movie* (A classroom action research at the eight grade of SMP Negeri 2 Grogol in 2013/2014 academic year). In Seminar Nasional Kajian Bahasa, Sastra, dan Pengajarannya (SEMNAS KBSP) IV 2016. <https://publikasiilmiah.ums.ac.id/xmlui/handle/11617/11282>
- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1-20). John Benjamins. <https://doi.org/10.1075/llt.32.01hou>
- 井長洋. (2019). 中学校外国語科の授業におけるリテリングの効果. 広島大学附属中・高等学校中等教育研究紀要, 66, 81-90.
- 岩井千秋. (2014). 英語によるオーラル・プレゼンテーションの指導効果の探索的検証: 学習者の言語産出面の変化に焦点を当てて. 大学英語教育学会中国・四国支部研究紀要, 11, 142-155.
- 岩男卓実. (2001). 文章生成における階層的な概念地図作成の効果. 教育心理学研究, 49(1), 11-20.
- Johnson, M. D., & Abdi Tabari, M. (2022). Task planning and oral L2 production: A research synthesis and meta-analysis. *Applied Linguistics*, 43(6), 1143-1164. <https://doi.org/10.1093/applin/amac026>
- Kawauchi, C. (2005). *Pre-task planning in L2 oral performance*. Kinseido.
- 茅野潤一郎. (2018). 即興的スピーキングに対する意識と学習経験—英語中級レベルの大学生の場合—. 中部地区英語教育学会紀要, 47, 17-24.
- Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, 4, 900-913. <https://doi.org/10.4304/jltr.4.5.900-913>
- Koizumi, R., In'nami, Y., & Jeon, E. H. (2022). L2 speaking and its internal correlates: A meta-analysis. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 proficiency: Theoretical and meta-analytic investigations* (pp. 307-338). John Benjamins. <https://doi.org/10.1075/bpa.13>
- 国立教育政策研究所. (2023). 『全国学力・学習状況調査の結果概要』 <https://www.nier.go.jp/23chousakekkahoukoku/report/data/23summary.pdf>
- 国立大学法人 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. 令和5年度「学力調査を活用した専門的な課題分析に関する調査研究」 https://www.mext.go.jp/content/20240517-mxt_chousa02-000036016_07.pdf
- Kormos, J. (2006). *Speech production and second language acquisition*. Routledge.
- Kurniatillah, R. E., Hidayat, D. N., Husna, N., & Alek, A. (2022). Teacher-student interaction in English classroom setting. *Journal of Applied Studies in Language*, 6(1), 53-63. <http://ojs2.pnb.ac.id/index.php/JASLhttp://ojs2.pnb.ac.id/index.php/JASL>
- 小柳かおる・峯布由紀 (2016). 『認知的アプローチから見た第二言語習得—日本語の文法習得と教室指導の効果』くろしお出版.
- Langacker, R. (1987). *Foundations of cognitive grammar: Vol. 1*. Stanford University Press.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417. <https://doi.org/10.1111/j.1467-1770.1990.tb00669>
- Levelt, W. (1989). *Speaking: From intention to articulation*. MIT press.
- McDonough, K., & Chaikitmongkol, W. (2010). Collaborative syntactic priming activities and EFL learners' production of wh-questions. *Canadian Modern Language Review*, 66(6), 817-841. <https://doi.org/10.3138/cmlr.66.6.817>
- 文部科学省. (2017a). 『中学校学習指導要領(平成29年告示)』 https://www.mext.go.jp/content/20230120-mxt_kyoiku02-100002604_02.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』開隆堂出版.
- 文部科学省. (2022). 高等学校外国語科におけるパフォーマンステスト参考資料(指導者用資料). https://www.mext.go.jp/content/20220715-mxt_kyoiku01-000021347_1.pdf
- 森下美和. (2017). インタクションはプライミングを引き起こすか: 自然な対話の中の疑問文に見る構文産出傾向. 言語学習と教育言語学: 2016 年度版, 85-90.
- 村端五郎, & 村端佳子. (2020). 用法基盤モデルの言語習得観にもとづいた小学校英語の展開. 小学校英語教育学会誌, 20(01), 148-

引用文献

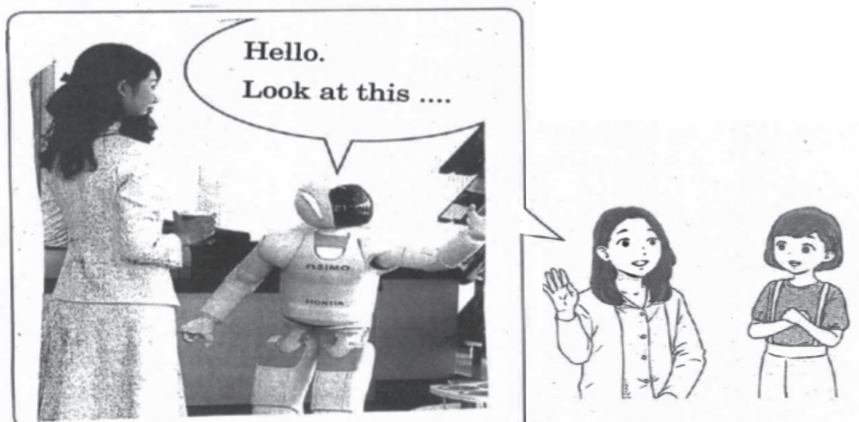
163. https://doi.org/10.20597/jesjournal.20.01_148
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21(1), 109-148. <https://doi.org/10.1017/S0272263199001047>
- Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 77-110). John Benjamins. <https://doi.org/10.1075/llt.11>
- Putri, O. H. Y., Perdhani, W. C., & Isnaini, M. H. (2021). The effect of life-story-retelling technique on student's speaking fluency. *The Journal of English Literacy Education: The Teaching and Learning of English as a Foreign Language*, 8(2), 113-124. <https://doi.org/10.36706/jele.v8i2.15260>
- 佐々木啓成. (2020). 『リテリングを活用した英語指導』大修館書店.
- Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language teaching research*, 1(3), 185-211. <https://doi.org/10.1177/136216889700100302>
- 田島充士. (2011). 再文脈化としての概念変化—ヴィゴツキー理論の観点から—. *心理学評論*, 54(3), 342-357. https://doi.org/10.24602/sjpr.54.3_342
- 高木哲也. (2022). 高校英語授業におけるリテリングの実践:効果的なリテリングを目指して. 筑波大学付属高等学校研究紀要, 63, 51-60.
- Tomasello, M. (2003). *Constructing a language: A usage-based-theory of language acquisition*. Harvard University Press.
- 富岡史子. (2016). 日本語 L1 幼児のノの意味のネットワーク形成から見るノダ. 岡山大学大学院社会文化科学 研究科紀要, 41, 127-141. <https://doi.org/10.18926/54204>
- Vercellotti, M. L. (2017). The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied linguistics*, 38(1), 90-111. <https://doi.org/10.1093/applin/amv002>
- Wendel, J. N. (1997). *Planning and second-language narrative production* (Publication No. 9813575) [Doctoral dissertation, Temple University]. ProQuest Dissertations and Theses.
- 米崎里, 松岡達也, & 米崎啓和. (2022). 小学校英語授業におけるリテリングの活用と効果. 中部地区英語教育学会紀要, 51, 157-164. https://doi.org/10.20713/celes.51.0_157
- Zuhriyah, M. (2017). Storytelling to improve students' speaking skill. *English Education: Jurnal Tadris Bahasa Inggris*, 10(1), 119-134. <https://doi.org/10.24042/ee-jtbi.v10i1.879>

資料1: プランニングで使用するメモ用紙(実験群)

Keyword Only

Class () Name ()

資料2: プランニングで使用するメモ用紙(比較群)



Class () Name ()

日本語母語英語学習者が使用する定型表現の分析

—「話すこと[やりとり]」と「話すこと[発表]」の技能育成をめざして—

研究者:小出 凱渡 茨城県／筑西市立明野中学校 教諭・茨城県／筑波大学大学院 在籍

《研究助言者:竹内 理》

概要

本調査では、定型表現の一種である高頻度語連鎖 (Lexical Bundles) を観点に、モノログとダイアログという話し言葉の2つの形態における日本語母語英語学習者の定型表現の使用実態を明らかにすることを目的とした。英語母語話者と学習者の話し言葉コーパスを比較し、学習者が過剰・過少使用する語連鎖を特定してリスト化するとともに、それらの語連鎖の特徴について分析・考察を行った。その結果、モノログでは、学習者は特定の句動詞を含む語連鎖や理由を述べる際に用いる表現を多用する傾向が見られた。一方、母語話者と比較して、学習者は多様な句動詞表現や助動詞を構成要素とする語連鎖をあまり使用していないことが明らかとなった。ダイアログでは、学習者は同じ発話の繰り返しで構成される語連鎖やフィラーを含む表現を過剰使用していた一方で、would like toや多様な時制表現を示す語連鎖を過少使用していることが明らかとなった。本調査の結果は、話し言葉における英語学習者の語連鎖の使用実態について有益な情報を提供し、定型表現を効果的に活用した「話すこと[やりとり]」及び「話すこと[発表]」の技能育成に示唆を与えるものである。

1 はじめに

近年、主にコーパスを活用した言語研究によって、話し言葉や書き言葉の多くが2語以上のまとまりからなる複数語で構成されていることが明らかにされてきた (Altenberg, 1998; Biber et al., 1999; Erman & Warren, 2000; Foster, 2001)。このように、自然談話内において一定の単語の組み合わせで繰り返し用いられる言語パターンは、定型表現 (formulaic language) と呼ばれており、外国語教育研究においても近年ますます注目を集めている (金澤, 2020)。定型表現は言語処理や言語習得と深く関わりがあり、正確で適切かつ流暢な言語使用に欠かせないものである (Nattinger & DeCarrico, 1992; Pawley & Syder, 1983; Schmitt & Carter, 2004; Wray, 2002)。定型表現は母語話者だけではなく、学習者にとってもその有用性が明らかにされており (Conklin & Schmitt, 2008; Isobe, 2011; Jiang & Nekrasova, 2007)、第二言語習得における定型表現の重要性が示唆されている。中田 (2022) は、英語学習者も定型表現の知識を身につけることで多くのメリットを享受することができるとし、言語使用における正確性・流暢性が向上することに加え、状況に応じた適切な言語使用ができるようになること、すでに知っている単語の知識が深まること、文法知識の習得が高まることなどを挙げている。

このように、定型表現は学習者にとって学習・習得の必要性が高いことは明らかであるものの、学習者にとって定型表現の習得・使用は必ずしも容易なものではない。先行研究では、学習者の使用する言語は母語話者と比べ定型性が低く (Granger, 1998; Howarth, 1998; Li & Schmitt, 2009)、学習者は、言語産

出の場面に応じて定型表現を適切に使い分けることができないことが報告されている (Chen & Barker, 2010, 2016)。このような問題を解決し、学習者が定型表現を効率的に習得するためには、学習者言語における定型表現使用の実態を詳細に調査し、その特徴を踏まえた指導・学習を行っていくことが重要である。

定型表現には、イディオムや句動詞、コロケーションなど複数のタイプが存在するが、その中の1つに高頻度語連鎖 (Lexical Bundles) がある。これまでの研究では、書き言葉における日本語母語英語学習者の語連鎖の使用 (石川, 2019) や日本の英語検定教科書に見られる語連鎖の扱いについて調査したもの (Mikajiri, 2022; Northbrook & Conklin, 2018) はいくつか存在するが、話し言葉における日本語母語英語学習者の語連鎖の使用実態を調査した研究は非常に限られている。また、現在の学習指導要領では、「話すこと」が [発表] と [やりとり] の2つの領域に分けられ、別々に目標や指導内容が定められている。これは、一人の話者が連続して話す場合と、複数の話者が相互に話す場合で、「話すこと」の特性が異なるためである (文部科学省, 2017b, 2018)。これらを踏まえ、本調査は、話し言葉を聞き手に対して一方向で話すモノローグ (Monologue) と、聞き手と話し手の双方向のやりとりであるダイアログ (Dialogue) の2つに区別し、コーパスを用いた先行研究の方法論に基づきながら、日本語を母語とする英語学習者の語連鎖使用の実態とその特徴を明らかにすることを目指した。

2 先行研究

2.1 言語使用における定型表現の役割

定型表現 (formulaic language) とは、複数の単語が組み合わせられ、1つのまとまりとして用いられる結束性の高い語の連鎖と定義され (Wray, 2008)、イディオムや句動詞、コロケーションなどが存在する。定型表現は、言語使用や言語学習において重要な役割を果たすと考えられている (Conklin & Schmitt, 2012; 金澤, 2020; 中田, 2022; Siyanova-Chanturia & Pellicer-Sánchez, 2019)。その理由として、第一に、定型表現の普遍性がある。Howarth (1998) では、およそ24万語の学術テキストを収集したコーパスを分析したところ、約30% から40% がコロケーション (e.g., blow the fuse) やイディオム (e.g., blow the guff) といった単語連鎖から構成されていることが明らかとなった。また、Erman and Warren (2000) によると、談話中に定型表現が占める割合は書き言葉で52.3%、話し言葉においては58.6% に及ぶとされている。Conklin and Schmitt (2012) では、先行研究で得られた知見を踏まえ、定型表現は会話の3分の1から半分程度を構成していると報告している。このように、自然談話内において定型表現は多くの割合を占め、言語を使用する上で必要不可欠なものであると考えられている。

第二に、定型表現は言語処理を効率化し、適切で流暢な発話を可能にするとされる。Jiang and Nekrasova (2007) は、英語母語話者とトルコ語や中国語、アラビア語など多様な言語的背景を持つ英語学習者を対象に文法性判断課題を実施したところ、定型表現 (e.g., to tell the truth) の方が、非定型表現 (e.g., to tell the price) よりも反応時間が短く、正答率が高いことがわかった。この結果から、母語話者や学習者にかかわらず、定型表現は1つの語彙項目として心的辞書に保有され、ひと固まりとして認識されることによって各構成語に注意を払う必要がなくなり、より短い時間で正確に言語処理がなされることが示唆された。Conklin and Schmitt (2008) は、英語母語話者と第二言語 (L2) として英語を学ぶ協力者を対象に、自己ペース読み課題を実施し、定型表現を含む文とそうでない文ではどちらが速く読解されるかについて検証した。分析の結果、母語話者と学習者ともに、定型表現を用いた文の方がより速く読解されたことから、定型表現は言語処理上の利点 (processing advantage) を持つと結論づけた。また、中田 (2022) は、定型表現の知識は、言語使用における流暢性 (fluency) を高め、正しい文を淀みなく流暢に産出したり、素早く内容を理解したりする上では不可欠であるとしている。さらに、定型表現を使用することで、話している相手や状況に応じた適切な言語使用が可能となり、The point is that や Help yourself

といった特定の機能を持つ定型表現を用いることで、効果的にメッセージを伝えたり、理解したりすることができるとして、言語使用における定型表現の重要性を指摘している。

以上のように、これまでに様々な研究から、言語使用における定型表現の重要性が示されてきている。

2.2 高頻度語連鎖 (Lexical Bundles) に着目した定型表現研究

定型表現と聞くと、イディオムや句動詞などをイメージすることが多いが、これまでの研究では、コーパスを活用し、その中に見られる高頻度語連鎖 (Lexical Bundles) を対象に、話者の使用する定型の言語表現の特徴を明らかにしようとしたものが多く行われている (e.g., Biber et al., 1999; Biber et al., 2004; Biber & Barbieri, 2007)。Lexical Bundles とは、ある談話内において、繰り返し用いられる結束性の高い語の連鎖である (Biber et al., 1999)。

Biber and Barbieri (2007) は、スピーキングとライティングといった2つの産出モードにおいて、英語母語話者が使用する高頻度語連鎖について分析を行った。この研究では、大学の様々な場面における発話を収集したコーパスを用い、100万語の中で40回以上出現する4語の高頻度語連鎖を対象に調査がなされた。話し言葉コーパスには、大学での講義場面や教員と学生による面談の場面などで発話されたデータが含まれている。一方、書き言葉コーパスは、講義で使用する複数の教科書や参考書、大学のパンフレットなどから抽出されたデータで構成されている。分析の結果、話し言葉と書き言葉のどの場面においても高頻度で使用する語連鎖が観察され、特に話し言葉では、教師が課題や授業の活動について指示する表現 (e.g., you just have to, you need to do), 学生に与えられた課題を遂行するように指示する表現 (e.g., you should be able to), 講義でトピックや活動について言及する際に用いる表現 (e.g., I want to talk about, what we're going to), 教材に注意を向けさせる表現 (e.g., if you look at) などが抽出された。それに対し、書き言葉では、時間を示す表現 (e.g., the beginning of each, the end of each, at the time of) に加え、講義の内容や構成について導入する表現 (e.g., an introduction to the, the nature of the) が観察された。

Jalilifar et al. (2016) は、芸術や文学、応用言語学、科学といった異なる学問分野に見られる高頻度語連鎖について調査を行った。それぞれの分野の研究論文を収集したコーパスを用いて100万回に10回以上出現する3, 4, 5語連鎖を抽出したところ、3語連鎖では、as well as や the number of, based on the, in terms of などが全ての分野に共通して出現していた。また、4語連鎖では、at the same time や in the case of, on the other hand, 5語連鎖では at the end of the や in the case of the, as a result of the といった表現が抽出された。特に、この研究で観察された4語連鎖の頻度順上位50位に含まれる項目の多くが、同様に複数の学問分野における Lexical Bundles を調査した Hyland (2008) で見られた語連鎖と類似していることが明らかとなった。

2.3 学習者における語連鎖の使用実態

2.2に挙げた研究はいずれも母語話者を対象としたものであるが、L2学習者を対象とした同様の調査も行われている。Chen and Baker (2010) では、中国語を母語とする英語学習者を対象に、学習者が使用する語連鎖 (4語連鎖) の言語的機能や言語的構造について調査を行った。学習者コーパスと母語話者コーパスを用いて分析を行ったところ、母語話者と比較して学習者は、the extent to which といった関係詞節の一部を構成する名詞句、the + 名詞 + of + the や in + the + 名詞 + of などの of を含む語連鎖をあまり使用しないことが明らかとなった。また、受動動詞 + 前置詞句を構成する語連鎖である be taken into account や be used in the に加え、自分の意見や主張を和らげる表現である is likely to be や would have to be などの表現も過少使用していた。一方で、学習者は母語話者と比べ、all over the world といった特定の表現や、不定詞節を導く語連鎖である in order to achieve や to be able to を多用していること

がわかった。さらに、学習者の語連鎖使用の特徴として、in the long runのように英語母語話者が主に話し言葉で使用する表現を書き言葉で用いる傾向が観察された。

日本語を母語とする英語学習者を対象とした研究としては、石川(2019)がある。石川は、複数のコーパスを用いて日本の英語教育で使用されている教材や試験問題における語連鎖、学習者の語連鎖使用の特徴について分析を行った。その結果、学習者が使用する語連鎖は、教材や試験問題に現れる語連鎖と比べて、母語話者が使用する語連鎖と大きく乖離しており、使用傾向が異なっていることが明らかとなった。また、学習者の過剰使用語連鎖の大半は、英作文の際に与えられたトピック文の一部となっており、トピック文に含まれる表現をそのまま流用する傾向が観察された。トピック文にない語連鎖としては、意見や主張を導くI thinkの過剰使用が特徴的であった。一方で、学習者が過少使用していた語連鎖を見ると、one of the やsome of the, as well asといった物事の量や比率、要素の添加などを表すものなどが抽出されており、母語話者と比較して学習者はこれらの表現をあまり使用しない傾向が見られた。

上記2つの研究のように、学習者の書き言葉における語連鎖使用の実態を調査した研究は他にもいくつか存在するが(e.g., Ådel & Erman, 2012; Güngör & Uysal, 2016; 小出, 2024; Pan et al., 2013; Staples et al., 2013), 話し言葉について分析した研究は限られている。Yan(2019)は、中国語を母語とする英語学習者を3つの習熟度に分け、話し言葉で使用する語連鎖の特徴について習熟度別に調査を行った。この研究では、Test for English Majors(TEM)と呼ばれる中国で実施されている英語力を測定する試験において、受験者が発話したデータを対象として分析を行った。その結果、習熟度が低い学習者は、量や数、程度の大きさを表す語連鎖(quantifying bundles)である more and more people やthere are a lotなどを多く使用する一方で、習熟度の高い学習者は、pay more attention toやintegral part of theといったトピックに焦点を当てる表現に加え、we should have aやwe should learn fromなどのshouldを含む語連鎖を多用することが示された。さらに、I would like toやis the most importantなどの自分の考えを表す語連鎖、話者の意見の確信度を示すI think it isやI don't think, トピックを導入する際に用いる表現であるwe all know thatやfirst of all Iなどは、3つの習熟度において同程度の使用頻度であり、話し言葉において学習者が共通して使用する語連鎖であることがわかった。特に、この研究で抽出されたon the other handやat the same timeなどのいくつかの表現は、学習者の書き言葉を調査した他の研究においても同様に観察されており(Chen & Baker, 2016; Staples et al., 2013), 産出モードの違いにかかわらず、学習者が好んで使用する語連鎖であることが明らかとなった。

3 研究の目的

先行研究では、日本語母語英語学習者を対象に、書き言葉における学習者と母語話者が使用する定型表現(語連鎖)の比較・分析がコーパスに基づいて行われ、学習者の定型表現の使用実態が調査されている(石川, 2019)。しかしながら、話し言葉について同様の調査はほとんど行われておらず、日本語母語英語学習者の定型表現の実態を明らかにする上で検証の余地が残る点である。そこで、本調査では、日本語母語英語学習者が話し言葉で使用する語連鎖に焦点を当て、先行研究と同様にコーパスを用いたアプローチによって、その実態と特徴を明らかにすることを目的とした。

現在の日本の学校英語教育において、「話すこと」が[発表]と[やりとり]の2つの領域に分けられて指導が行われていることを踏まえ、本調査では、話し言葉を聞き手に対して一方向で話すモノローグ(Monologue)と、聞き手と話し手の双方向のやりとりであるダイアログ(Dialogue)の2つに区別し、それぞれの産出モードにおける学習者の定型的言語表現の分析を行うこととした。研究課題は以下の2つである。

RQ1

モノローグ (Monologue) において、日本語母語英語学習者が過剰・過少使用する語連鎖は、どのようなものがあるか。また、その特徴 (品詞、文法、語用論的機能など) は何か。

RQ2

ダイアログ (Dialogue) において、日本語母語英語学習者が過剰・過少使用する語連鎖は、どのようなものがあるか。また、その特徴 (品詞、文法、語用論的機能など) は何か。

4 研究の方法

本節では、調査に用いる各コーパスの基本情報、分析対象とする語連鎖の基準、そしてコーパス分析から得られたデータの分析手法について詳述する。

4.1 対象とするコーパスデータ

4.1.1 ICNALE の概要

本調査では、ICNALE (International Corpus Network of Asian Learners of English) コーパスにおける発話データを分析対象とした。ICNALE は、アジア圏における10の国や地域の英語学習者 (大学 (院) 生) と英語母語話者およそ4,000人による発話や英作文データを収集した大規模コーパスであり、14,000以上のサンプル数で構成されている (Ishikawa, 2023)。参加者の習熟度を判定するため、全ての学習者には、Common European Framework of Reference (CEFR) の基準に基づき、A2, B1_1, B1_2, B2の4段階で習熟度情報が付与されている。ICNALE は5つのモジュールから構成されており、1分間の独話データを集めた Spoken Monologue (SM)、インタビュー形式の対話を集めた Spoken Dialogue (SD)、英作文データを収集した Written Essays (WE)、学生の作文を専門家が校閲した Edited Essays (EE)、学生の作文・発話の評価データを収集した Global Rating Archive が含まれている (Ishikawa, 2023)。

本調査で ICNALE を分析対象とした理由は、以下の2つである。1つ目は、言語の産出条件が統制されている点である。ICNALE の全モジュールにおいて、学習者と母語話者は、「大学生のアルバイトの是非 (It is important for college students to have a part-time job.)」及び「レストラン全面禁煙の是非 (Smoking should be completely banned at all the restaurants in the country.)」という2つのトピックについて作文もしくは発話をし、データが収集されている。また、詳細は後述するが、SM や SD では、全ての参加者が同一条件の下で言語産出ができるように、事前に定められた順序に従いながらデータ収集がなされている。このような ICNALE の特性は、母語話者と学習者での比較や異なる母語を持つ学習者間での比較、習熟度別による比較といった対照分析の妥当性を高め、本調査のように、母語話者と比較して学習者言語の特徴を解明しようとする際に適したコーパスであるといえる。

2つ目は、2種の話し言葉モジュールが含まれている点である。ICNALE の話し言葉データには、話者が単方向で話した独話データである SM と、双方向のやりとりを集めた対話データである SD の2つがある。これらは、本調査で示唆を与えたいと考える「話すこと [発表]」と「話すこと [やりとり]」の特性を反映したものであると考える。以下、これら2つのモジュールにおけるデータを Ishikawa (2023) と石川ほか (2020) に基づいてさらに詳しく述べる。

4.1.2 ICNALE Spoken Monologue

ICNALE Spoken Monologue は、1,100人の参加者に、上述の「大学生のアルバイトの是非」と「レストラン全面禁煙の是非」という2つのテーマについて、それぞれ2回、1分間のスピーチを行わせた独話データを含む (Ishikawa, 2023)。まず、参加者は指定された電話番号に電話をかけ、システムから自動で流れる指示に従い、自己紹介発話 (60秒) を実施する。次に、1つ目のテーマである「大学生のアルバイトの是非」

について、第1発話(20秒準備後に60秒で実施)を行い、その後、第2発話(10秒準備後に60秒で実施)をする。ここで2度発話を行う理由は、緊張により第1発話で本来のパフォーマンスが発揮できない参加者がいる可能性を考慮したためである。自己紹介の発話はコーパスに含まれず、第1発話と第2発話がコーパスに含まれている。最後に、2つ目のトピックである「レストラン全面禁煙の是非」について、「大学生のアルバイトの是非」の発話手順と同様に、第1発話、第2発話を連続して行う。全ての参加者は同一条件の下、発話を行っており、発話データを収集後、それらの音声の書き起こしテキストを構築している。本調査では、その中の日本語母語英語学習者($n = 150$)と英語母語話者($n = 150$)の全データ(第1発話と第2発話)を調査対象とする。

4.1.3 ICNALE Spoken Dialogue

ICNALE Spoken Dialogue は、参加者425名に対してインタビューを行い、その発話を書き起こしたデータである(石川ほか, 2020)。インタビューの進め方は詳細なプロトコルブックによって定められており、①導入対話→②「大学生アルバイト」関連タスク→③「レストラン禁煙」関連タスク→④振り返りの手順で進み、計4部で構成されている。①は、インタビューの導入部分であり、英語学習などをテーマとした質疑が行われる。

②と③では、それぞれ、ICNALEの共通テーマである「大学生のアルバイト」と「レストランでの禁煙」に関連した課題が設定されており、(a)イラスト描写、(b)イラスト関連質疑、(c)ロールプレイ、(d)ロールプレイ関連質疑の4つのタスクで構成されている。③を例にすると、(a)のイラスト描写では、公園でタバコを吸っている男性に対して、子連れ母親が注意して公園内での喫煙を止めさせるという内容の6コマ漫画を時系列で描写させるというタスクが与えられる。(b)では、漫画の内容に関連した質問(公園に行くのは好きか、好きな公園はどこかなど)を提示して、参加者に回答させる。(c)では、インタビュワーと参加者がそれぞれ、レストランの店主役とその客役を演じてロールプレイを行う。参加者は、レストランでのタバコの臭いに耐えられず店を出ざるを得なかったことを抗議し、店主に返金を求める客として、インタビュワーを説得するという課題が与えられる。(d)では、ロールプレイの内容に関連した質問(どんなレストランが好きか、レストランの全面禁煙についてどう思うかなど)について参加者に答えさせる。最後に、④では、(a)英語による振り返り質疑と、(b)母語による振り返り質疑を続けて行わせる。なお、SDの日本モジュールに関しては、1名の非英語母語話者の教員が100名全てのインタビューを行っており、これにより、調査者の性別や外国語知識などといった違いによるインタビューへの影響が抑制され、データの統制性の向上にも寄与すると考えられる(石川, 2018)。

SDデータのうち、本調査では、日本語母語英語学習者($n = 100$)と英語母語話者($n = 20$)の発話データを調査対象とする。また、本調査では、関連タスクのうち、(b)イラスト関連質疑、(c)ロールプレイ、(d)ロールプレイ関連質疑の計6部のデータのみを分析対象とし、②と③の(a)イラスト描写と、①導入対話と④振り返りの4つのパートを分析対象から除外した。これらのデータを分析対象外とした理由として、①と④は、インタビューの導入や振り返りのパートであり、実際に2つのテーマについてのタスクを行っている発話ではないこと。また、②と③の(a)は、与えられた課題についての発話ではあるが、参加者が一方的にイラストについて描写しており、モノログの特性に近いデータであると判断したためである。以上2種のコーパスの基本情報に関しては、表1を参照されたい。

■表1: 本調査で分析対象とした各コーパスの基本情報

	Spoken Monologue	Spoken Dialogue
形式	独話型	対話型
参加者	日本語母語英語学習者 150名	日本語母語英語学習者 100名
	英語母語話者 150名	英語母語話者 20名
トピック	大学生のアルバイト レストランでの禁煙	大学生のアルバイト レストランでの禁煙
タスク	意見主張	イラスト関連質疑 ロールプレイ ロールプレイ関連質疑
総語数	英語母語話者 94,168	英語母語話者 30,753
	日本語母語英語学習者 41,737	日本語母語英語学習者 103,073
総ファイル数	英語母語話者 600	英語母語話者 120
	日本語母語英語学習者 600	日本語母語英語学習者 600

4.2 語連鎖の抽出基準

高頻度語連鎖を対象とした多くの研究においては、任意の語連鎖が定型表現であるかどうか、すなわち、話者が高頻度で繰り返し用いる表現であるかどうかについて判定する方法として、頻度 (frequency) と分布度 (dispersion) の2つの指標が用いられている (Ädel & Erman, 2012; Biber et al., 2004; Biber & Barbieri, 2007; Chen & Baker, 2016; Hyland, 2008; Yan, 2019; 石川, 2019)。頻度については、研究者によって様々な閾値が用いられている (Biber et al., 2004; Hyland, 2008)。英語学習者の話し言葉における語連鎖を調査した Yan (2019) では、同様に英語母語話者の話し言葉について検証した Biber et al. を参考に、100万語中40回以上の頻度で現れる語連鎖を抽出対象としている。本調査では、話し言葉を分析対象とすることを鑑み、話し言葉における語連鎖を調査した Yan と Biber et al. を参考に、100万語中に40回以上出現する語連鎖を対象にした。なお、Yan と同様、本研究で使用した4つのコーパスはどれも100万語を満たしていないため、100万語に換算したときに40回以上となるものを抽出したことに注意されたい。

上述の通り、頻度の抽出基準については研究によって様々であり、どの程度の頻度であれば高頻度としてみなされるかについては明らかではないため、頻度のみを用いて定型表現を判定するのは注意を要する。そこで重要となる2つ目の指標が、分布度である。Biber et al. (2004) では、高頻度語連鎖は、コーパス内のテキストに幅広く分布していると主張し、分布度を用いることで個々人の偏った語連鎖使用で見られる表現の抽出を防ぐことができるとして、高頻度語連鎖を異なる5つ以上のファイルに用いられるものと

定めている。これを踏まえ、Yan(2019)では、分析対象とした3つのコーパスの総ファイル数がそれぞれ276, 314, 217と異なっていることを考慮し、また、Biber et al. が用いた5と近接するように、各コーパスの総ファイル数のうち2%以上のファイルに出現する語連鎖を抽出している。上記2つの先行研究を参考に、本調査においても対象とするコーパスサイズが異なること、分布度が先行研究に近似した値(5)になることに留意し、各コーパスの総ファイル数のうち1%以上のテキストに出現する語連鎖を対象とすることとした。

最後に、どの長さの語連鎖を抽出するかについてであるが、これまでに最も多く研究されてきた長さは4語であるとされている(Chen & Baker, 2010)。その理由として、4語連鎖は、手動で語連鎖を分類したり、コンコーダンスを確認したりする際に最も扱いやすく(Chen & Baker, 2010)、5語連鎖よりも頻度が高く現れ、3語連鎖よりもはっきりとした様々な構造や機能を示すとされるためである(Hyland, 2008)。また、話し言葉を調査したBiber et al.(2004)やYan(2019)でも4語連鎖を扱っており、4語を対象とすることで、多くの先行研究との整合性を図ることができる。これらを踏まえ、本調査においても4語連鎖について分析することとした。

表2は、これまでに述べたコーパスごとの抽出基準を示したものである。ただし、ICNALE Spoken Monologueの英語母語話者と日本語母語英語学習者、ICNALE Spoken Dialogueの日本語母語英語学習者については、分布度との関連で、頻度の抽出基準は実質的には6以上となっていることに留意されたい。

■表2: 各コーパスにおける語連鎖の抽出基準

コーパス	頻度 (frequency)		分布度 (dispersion)	
	総語数	抽出基準	総ファイル数	抽出基準
ICNALE Spoken Monologue (英語母語話者)	94,168	4	600	6
ICNALE Spoken Monologue (日本語母語英語学習者)	41,737	2	600	6
ICNALE Spoken Dialogue (英語母語話者)	30,753	2	120	2
ICNALE Spoken Dialogue (日本語母語英語学習者)	103,073	5	600	6

4.3 語連鎖の抽出方法

語連鎖の抽出は、フリーコーパスソフトウェアであるAntConc(Anthony, 2023)のN-gram機能を用いて実施した。N-gramは、コーパス内に共通して現れる任意の長さNの語連鎖を抽出する分析機能であり、Nの語連鎖とは、it isのような2語連鎖やas well asのような3語連鎖を示す。なお、本調査では書き言葉における日本語母語英語学習者の語連鎖使用について調査した石川(2019)に倣い、抽出した各項目に動詞や形容詞の活用形が含まれる場合も、原形に戻さず、テキストの中で用いられているままの表記で集計を行った(例: is going toとam going toは別個に処理)。また、AntConcでの処理の際、抽出された語連鎖は、全て小文字扱いとなり(例: I want toは、i want toとして処理)、ハイフン、アポストロフィ、コンマ、ピリオド、クエスチョンマークといった記号は1語と見なさず、スペース扱いとなり省略され処理されている(例: part-timeはpartとtimeの2語、can'tは、canとtの2語として処理)。

4.4 トピックに関連した語連鎖の扱い

石川(2019)によると、学習者は言語産出の際に与えられた指示文に含まれる表現をそのまま流用する傾向が見られ、学習者が過剰使用する語連鎖の大半はトピックに関連した表現であるとされる。この点を踏まえ、本調査では、トピック文に含まれる表現や直接関連した表現以外の高頻度語連鎖に焦点を当てるため、語連鎖の抽出後、トピックに関連した語連鎖(topic bundles)を分析対象から除外した。具体的には、a part time jobやsmoking is bad forといったICNALEの共通トピックに関連する語連鎖、SDの絵描写やロールプレイの内容に関する表現であるswim in the sea やsmoking in the parkなどは分析対象から除外した(絵描写の発話自体は分析に含めていないが、このような表現は主に関連質疑において見られた)。このような表現からは、定型表現に関する教育的示唆を得られる可能性が低く、話し言葉を調査したYan(2019)においても分析対象外とされていることから、本調査でも対象外とした。

4.5 語連鎖の分析方法

本調査では、学習者が使用する語連鎖の使用実態を調査するため、日本語母語英語学習者の書き言葉を対象とした石川(2019)を参考に分析を行うこととした。分析の手順は以下の通りである。

初めに、各語連鎖の1万語あたりの調整頻度と分布度比率を算出した。ここで、調整頻度を1万語あたりとした理由は、元のコーパスサイズを超えて基準値を設定することは、データが持つ情報量を過大解釈することにつながるためである(石川, 2021)。したがって、基準値の決定については、調整しようとする全てのコーパスサイズの下限値を超えない範囲で最も大きな10の乗数を選択するのが原則であるとされる(石川, 2021)。そのため、本調査で分析対象とした4つのコーパスサイズを考慮し、1万語あたりの調整頻度を用いることとした。

次に、分布度比率の算出を行った。分布度比率とは、各コーパスに所収の総ファイル数のうち、当該語連鎖が1回以上出現しているファイル数の比率である(例: 総ファイル数が300であり、そのうち任意の語連鎖が1回以上出現しているファイル数が210の場合、分布度比率は70%となる)。

最後に、1万語あたりの調整頻度に分布度比率を掛けた値である重要度を求めた。重要度は、頻度と分布度を同時に評価する指標であり、例えば、調整頻度が200で分布度比率が50%の語連鎖と、調整頻度が150で分布度比率が80%の語連鎖の場合、前者の重要度指標は100、後者の重要度指標は120となる(石川, 2019)。

その後、各語連鎖における母語話者の重要度から学習者の重要度を引いた差分値を算出した。重要度の値が正で高いほど学習者が過少使用している表現であり、負で低いほど学習者が過剰に使用している語連鎖であると解釈される。ただし、学習者データもしくは母語話者データ的一方のみから抽出された項目については、抽出されていない方の当該語連鎖の重要度を0として差分値を求めた。

差分値を算出した後、2つのRQについて順に調査を行った。まず、RQ1について、SMにおいて学習者が過剰・過少使用している差分値上位100位までの項目を取り上げ、過剰・過少使用のそれぞれについて差分値の絶対値順にリスト化した。次に、作成したリストを元に、学習者が使用する高頻度語連鎖の言語的特徴について、各語連鎖を構成する語の品詞やその表現の文法について検討した。また、抽出した語連鎖の前後の文脈についても観察し、その語用論的機能について検討した。RQ2については、SDを対象に同様の手順で分析を行った。

5 結果と考察

表3は、SMとSDそれぞれで抽出された語連鎖数である。これらの語連鎖のうち、本調査では4.4で述べ

た通り, non-topic bundlesを対象に分析を行った。また, 表4と表5は抽出された語連鎖について, 語の品詞や文法事項に着目し, いくつかのカテゴリーに分類したものである。表の左側には各カテゴリーの名称, 中央には本研究で抽出された語連鎖のうち, そのカテゴリーに分類された特徴的な語連鎖, 右側には各語連鎖の差分値が記載されている。これらのカテゴリーごとに, 本節では, 学習者の語連鎖使用の実態をその考察と併せて報告する。

■表3: 各コーパスから抽出した語連鎖数

カテゴリー	抽出した語連鎖		
コーパス	all bundles	topic bundles	non-topic bundles
ICNALE Spoken Monologue (英語母語話者)	455	203	252
ICNALE Spoken Monologue (日本語母語英語学習者)	263	126	137
ICNALE Spoken Dialogue (英語母語話者)	730	104	626
ICNALE Spoken Dialogue (日本語母語英語学習者)	541	115	426

■表4: SMにおいて抽出した語連鎖のカテゴリー化

カテゴリー	項目	差分値
句動詞を含む語連鎖	i agree with this	-4.1131
	it depends on the	0.0234
意見主張を示す語連鎖	i think it is	-0.6074
	believe that it is	0.0299
助動詞を含む語連鎖	i think we should	-0.1138
	should be able to	0.0673
その他の語連鎖	i have two reasons	-0.3358
	at the same time	0.1168

■表5: SDにおいて抽出した語連鎖のカテゴリー化

カテゴリー	項目	差分値
願望を表す語連鎖	so i want to	-0.3936
	i would like to	0.2212
発話の繰り返しやフィラーを 構成語彙とする語連鎖	i can i can	-0.2411
	uh let me see	0.0244
その他の語連鎖	i am going to	0.2385
	as long as they	0.0542

5.1 モノローグ (Monologue) における学習者の語連鎖使用の実態 (RQ1)

5.1.1 句動詞を含む語連鎖

過剰使用語連鎖の差分値順上位100位のうち、agreeを含む語連鎖が23項目、disagreeを含む語連鎖は5項目観察された。このうち、句動詞であるagree withやdisagree withが含まれていた項目は計19項目であった。特に、i agree with thisは、過剰使用語連鎖の中で最も差分値の高い項目であり、母語話者に比べ学習者がとりわけ多用する表現であるといえる。また、i disagree with thisにおいても差分値順位が5位であったことから、母語話者と比較し学習者が好んで使用する表現であると考えられる。書き言葉を調査した石川(2019)においても、i agree with thisとi disagree with thisは共に学習者が過剰使用する表現として抽出されており、これらは話し言葉や書き言葉にかかわらず、母語話者と比べ学習者が多用する語連鎖であるといえるだろう。さらに、agreeを構成語彙とする語連鎖の中には、i agree this statementやi agree the opinionなどのagree＋目的語の構成を持つ語連鎖が9項目観察された。これらはwithが不足した文法的に誤った表現であり、学習者コーパスからのみ抽出され、母語話者は使用しない表現であった。

一方、学習者が過少使用する語連鎖に注目すると、have to worry aboutやit depends on the, to take care ofなどの句動詞を含む語連鎖が抽出され、これらの学習者の使用頻度は0であり、母語話者のみが使用した表現であることがわかった。話し言葉における句動詞使用を調査した石井(2018)では、句動詞使用について、学習者の習熟度が上がるにつれて使用頻度は上昇していくものの、上級の学習者であっても使用する句動詞は母語話者とは大きく異なり、母語話者の使用頻度が高いにもかかわらず、学習者がうまく使えないものがあることを報告している。本調査においてもその傾向が現れ、上記のような句動詞を含む語連鎖の使用傾向に違いが見られたと考えられる。

5.1.2 意見主張を示す語連鎖

これまでの研究では、書き言葉において、日本語母語英語学習者のi thinkのようなthinkを構成語彙とする語連鎖の過剰使用が報告されていた(石川, 2019; 阪上, 2013)。話し言葉を対象とした本調査においても、過剰使用語連鎖上位100位を見ると、10項目がthinkを含む語連鎖であった(e.g., i think it is)。一方、過少使用語連鎖では24項目が抽出されており(e.g., think it s important)、話し言葉(SM)においては、学習者だけではなく、母語話者についてもthinkを構成語彙とする語連鎖を多用する傾向があるといえる。

さらに、学習者の過少使用では、believe that it isやi believe it sのようなbelieveを含む表現が抽出されており、母語話者は自分の考えや意見を述べる際に、thinkに加え、believeを用いていると考えられる。加えて、学習者は、i do think thatやi do believe thatといった強調の助動詞doを構成語彙とする語連鎖を過少使用していた。

5.1.3 助動詞を含む語連鎖

次に、助動詞を含む語連鎖に注目すると、過剰使用語連鎖では、i think we shouldのみが抽出された一方で、過少使用では、計13項目が見られた。これらの表現の構造を確認すると、i think they shouldやi think it wouldなどのthink＋主語＋助動詞の構造を持つ語連鎖に加え、should not be allowedのように受動動詞を持つ表現やshould be able toといった3語連鎖であるbe able toを含む項目が見られ、学習者に比べ母語話者は、助動詞を含む多様な構造の語連鎖を使用していることがわかった。助動詞の使用について、これまでの研究で学習者と母語話者では助動詞の使用傾向が異なることが報告されている(石田, 2011; 鈴木, 2022)。特に、wouldについては、そもそも学習者は母語話者に比べ、wouldをあまり使用しないことが知られている(Ishikawa, 2019; 鈴木, 2022)。同様に、本調査でも学習者においてwouldを含む語連鎖の過少使用が見られたことになる。

5.1.4 その他の語連鎖

ここでは、差分値上位に占める項目数は少ないが、特徴的な構造や機能を持つ語連鎖を記述する。学習者の過少使用語連鎖では、it s going toやthey re going toといったbe going toを含む表現に加え、be able to haveやnot be able toなどのbe able toを構成語彙とする表現、when i was a やas i was saying, while they are in, if you don tといった多様な副詞節を導く語連鎖が抽出された。また、母語話者と比較して学習者はat the same timeやon the other handを過少使用していた。中国語母語英語学習者の書き言葉を調査したChen and Baker (2016)や話し言葉を扱ったYan (2019)では、at the same timeやon the other handは学習者コーパスから抽出されたが、本調査においては確認されなかった。このことから、これらは日本語母語英語学習者が特徴的にあまり使用しない語連鎖であると推察される。上記に挙げた表現については、学習者にとって習得が困難であり、定着に時間のかかる項目である可能性がある。

次に、過剰使用語連鎖を見ると、i have two reasons, first reason is thatなどのreasonを含む表現が4項目観察された。これらは、母語話者は一切使用せず、学習者のみが使用した語連鎖であった。これらはスピーチなどを行う際に典型的に指導される表現であると推察されるが、そのような表現が学習者にある程度定着しているといえる一方、母語話者がほとんど用いることのない表現を繰り返し指導していることにもなり得る。

5.2 ダイアログ (Dialogue) における学習者の語連鎖使用の実態 (RQ2)

5.2.1 願望を表す語連鎖

過剰使用語連鎖の差分値上位100位のうち、wantを含む語連鎖は28項目であり、およそ3割を占めていた。これらの語連鎖の構造を確認すると、so i want toやi want to goなどの不定詞節を導く表現や、i want you toのようにwant＋目的語＋toの構造を持つ項目が散見された。一方、過少使用語連鎖を見ると、want toと類似した意味を持つwould like toを含む項目が抽出され、願望を表す語連鎖として学習者がwant toを多用する一方で、母語話者はwould like toを用いて表現していると考えられる。

5.2.2 発話の繰り返しやフィラーを構成語彙とする語連鎖

過剰使用語連鎖の中には、when i when iやwe can we canのように同じ発話の繰り返しで構成されている語連鎖が9項目見られており、SMでは5項目抽出されていた。過少使用語連鎖においてもSDでこのような項目は5項目観察されたが、SMでは1項目のみであった。この結果から、学習者と母語話者共に繰り返しの発話を含む語連鎖をある程度使用すると考えられるが、比較的学习者の方がその傾向が強く、SMよりもSDにおいてその使用頻度は上昇するといえる。

また、SMの差分値上位の語連鎖では確認されなかったが、SDの過剰使用語連鎖では、umm i want toやuh i don tといったフィラーを含む表現が22項目存在していた。上記のような語連鎖が見られた理由については、学習対象の言語産出に伴う認知的負荷が高くなることで同じ語連鎖を繰り返し使用したり、言語産出に時間を確保するために、フィラーを含む表現を多用したりするためであると考えられる。特に、自分のペースで単方向的に話すSMと違って、SDは他者との対話であり、相手を意識した状況となるため、このような表現の使用頻度が増加する傾向があると考えられる。

加えて、SDでは、yes yes yes yesのような構成語彙が全て同じ項目が2つ観察された。これらの表現はSMでは抽出されなかったが、SDは対話データでありインタビューが質問し、参加者が回答する場面が多く存在しているため、上記のような語連鎖が見られたといえるだろう。ここで過少使用語連鎖を見ると、yes i do iやno i don tが抽出されていることから、回答する際に用いる表現についても学習者と母語話者では違いがあると考えられる。さらに、学習者はuh let me seeやwell i think itなどの語連鎖をあまり使用しておらず、会話をつなぐ表現の1つであるlet me seeやwellを含む表現を過少使用していた。

5.2.3 その他の語連鎖

過少使用語連鎖では、i was working inやi am trying toのような過去進行形や現在進行形、現在完了形を示すi ve been toや過去の習慣を表現するi used to go、予定を表すi am not goingなど多様な時制表現を含む語連鎖が見られた。また、学習者はas long as theyやas you can seeといったasを構成語彙とする表現に加え、SMでも過少使用として抽出されたat the same timeをあまり使用していなかった。上記に挙げた語連鎖については、学習者が時制表現の用法や意味をあまり理解できていなかったことや、理解できていたとしてもアウトプットする練習があまり与えられていなかったことから、実際に使用できなかった可能性が考えられる。

6 結論と今後の課題

本調査では2つの研究課題を通して、話し言葉における日本語母語英語学習者の語連鎖使用の実態を調査した。RQ1では、モノログを対象に分析を行い、母語話者と比較して学習者が過剰使用する語連鎖やあまり使用していない表現などについて検証した。その結果、学習者は特定の句動詞を含む語連鎖 (e.g., agree with) や理由を述べる際に用いる語連鎖 (e.g., i have two reasons) を多用する傾向が見られた。一方で、学習者の過少使用を見ると、他の多様な句動詞表現 (e.g., worry about, depends on) や助動詞を構成語彙とする表現に加え (e.g., i think we should, should not be allowed), at the same timeやon the other handなどの定型表現が抽出され、学習者は母語話者と比較してこれらの表現をあまり使用していないことが明らかとなった。また、自身の意見や考えを示す定型表現に着目すると、学習者と母語話者共にthinkを含む語連鎖を使用しており、さらに学習者はbelieve that it isのようにbelieveを構成語彙とする語連鎖を過少使用していることが明らかとなった。

次に、RQ2では、同様の手順でダイアログについて分析を行った。対話データにおける学習者言語の特徴として、学習者は同じ発話の繰り返しで構成される語連鎖やフィラーを含む項目 (e.g., when i when i, uh i don t), wantを構成語彙とする語連鎖 (e.g., so i want to, i want to go) を過剰使用していた一方で、would like toや多様な時制表現を示す表現 (e.g., i was working in, i ve been to), asを含む語連鎖 (e.g., as long as they) などを過少使用することが明らかとなった。フィラーや繰り返しを含む語連鎖を多く使用するという傾向は、モノログとは異なるダイアログの特性を反映しているといえる。

本調査で示した語連鎖のデータは、同一のテーマ・条件で収集された母語話者の発話と比較して、学習者が過剰使用・過少使用する表現を抽出したものであり、妥当性が高く、一定の資料的価値があると考えられる。本調査で得られた結果を基に、発話における学習者言語の特徴を把握することで、学習者がよく使う表現と類似した別の言い方や、会話場面で方略的に使われる表現などを指導したり、特定の動詞とよく使われる前置詞や副詞を組み合わせる指導したりするなど、指導上の工夫を考えることができるだろう。本調査の結果が、定型表現を効果的に活用した「話すこと[やりとり]」及び「話すこと[発表]」の技能育成のための一助となることを期待する。

一方で、本調査にはいくつかの限界点が存在することを述べておく必要がある。まず、本調査で利用したデータについて、SMは2014年頃、SDは2018年頃の時点で収集されたものであり、現在の日本国内における英語学習者の現状を反映したものではない点に留意が必要である。同様に、本調査で分析対象としたSMとSDは、現行の学習指導要領で示されている「話すこと[やりとり]」と「話すこと[発表]」の特性を厳密に反映したものではない。例えば、やりとりでは、即興で互いの考えや気持ちなどを伝え合ったりすることに加え、相手の発話に応じたり、それに関連した質問や意見を述べたりして、互いに協力して対話を継続・発展させることが必要とされているが(学習指導要領, 2017b)、SDではインタビューする人とされる人という役割が与えられ、主に事前に決められた順序で会話が進むという構成になっていた。また、発表では、スポーツや音楽など身の回りのことで生徒が共通して関心を持っていることについての話す

力の育成が示されているが、SMでは指定された2つのトピックのみについての発話であった。今後は、現在あるいは近年の学校英語教育を経験した学習者からデータを収集したり、「話すこと[やりとり]」と「話すこと[発表]」の特性をより厳密に反映したデータを用いて検証を重ねていくことが求められるだろう。

また、本研究で分析対象としたコーパスは、大学(院)生が産出した言語データであったため、中学生や高校生といった発達段階における学習者の語連鎖の使用については調査できていない。今後、他の発達段階における学習者の言語データも分析することで、各発達段階の語連鎖使用の実態についても明らかにすることができるだろう。

さらに、本研究では英語母語話者の定型表現の使用傾向を基準に、学習者の過剰・過少使用する表現を求めているが、母語話者の言語使用を厳密な基準とすることには議論の余地があるだろう。近年、批判的応用言語学の観点から、英語教育を取り巻く規範とその問題点の検討がなされている(久保田, 2018)。この批判的応用言語学の視点から見ると、母語話者を基準にした教育を行うことは現実と乖離しているとも考えられる。また、このような教育は、ネイティブスピーカリズム(native-speakerism)を助長する可能性もある。これらの点を考慮した上で、本研究の結果をどのように言語指導や学習に取り入れていくかを慎重に検討していく必要があるだろう。

最後に、本研究で分析した学習者と母語話者コーパスは総語数やサンプル数に大きな違いがあり、加えて、学習者コーパスでは習熟度に偏りがあることに留意が必要である。これらの要素は本研究で得られた結果に影響を及ぼしている可能性があり、その点を踏まえた上で本研究の結果は解釈される必要がある。今後はデータの均衡を図った上で、学習者の習熟度別に語連鎖使用の特徴などを検討することができれば、定型表現の学習・指導に対してさらに有益な示唆が得られるであろう。

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。とりわけ、私の調査を担当してくださり、激励のお言葉や丁寧なご指導をいただいた研究助言者の竹内理先生に厚く感謝申し上げます。そして、筑波大学大学院の名畑目真吾先生には、研究の着想から、調査、論文執筆まで親身なご指導をいただきました。心より感謝申し上げます。最後に、多くのご助言と励ましをいただきました研究室の皆様に深く御礼申し上げます。今後もこの貴重な経験を生かして、教育現場に少しでも還元できるよう、研究に精進いたします。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

- Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31(2), 81-92.
- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp.101-122). Oxford University Press.
- Anthony, L. (2023). AntConc (Version4.2.0) [Computer software]. Tokyo: Waseda University. Retrieved from <https://www.laurenceanthony.net/software/antconc/>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of Spoken and Written English*. Longman.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263-86.
- Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, 14, 30-49.
- Chen, Y., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, 37(6), 849-880.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? . *Applied linguistics*, 29(1), 72-89.
- Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45-61.

引用文献

- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29-62.
- Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning. Teaching and testing* (pp. 75-93). Longman.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp.79-100). Oxford University Press.
- Güngör, F. & Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native scholars. *English Language Teaching*, 9(6), 176-188.
- Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp.155-175). Oxford University Press.
- Hyland, K. (2008). As can be seen: lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4-21.
- 石井康毅 (2018).「話し言葉コーパスと検定教科書に基づく日本人英語学習者の句動詞使用実態の分析」*Learner Corpus Studies in Asia and the World*, 3, 101-119.
- 石川慎一郎 (2018).「The ICNALE Spoken Dialogue の設計:対話における L2口頭産出研究のために」*Learner corpus studies in Asia and the world*, 3, 9-26.
- 石川慎一郎 (2019).「英語教育における連語:ターゲット・インプット・アウトプットの三元コーパス分析をふまえた English N-gram List for Japanese Learners of English (ENL-J) の開発と利用」仁科恭徳・吉村由佳・吉川祐介 (編).『言語分析のフロンティア』(pp.32-47) 金星堂.
- 石川慎一郎 (2020).「第4章 コーパスと英語教育研究」石川慎一郎・長谷部陽一郎・住吉誠『コーパス研究の展望』(pp. 193-244). 開拓社.
- 石川慎一郎 (2021).『ベーシックコーパス言語学 第2版』ひつじ書房.
- Ishikawa, S. (2019). The ICNALE Spoken Dialogue: A New Dataset for the Study of Asian Learners' Performance in L2 English Interviews. *English Teaching*, 74(4), 153-177.
- Ishikawa, S. (2023). *The ICNALE Guide: An Introduction to a Learner Corpus Study on Asian Learners' L2 English*. Routledge.
- 石田知美 (2011).「日本人英語学習者コーパスを用いた法助動詞の使用に関する研究」『JACET 中部支部紀要』第9号, 57-73.
- Isobe, Y. (2011). Representation and processing of formulaic sequences in L2 mental lexicon: How do Japanese EFL learners process multi-word expressions. *JACET Kansai Journal*, 13, 38-49
- Jalilifar, A., & Ghoreishi, S. M., & Roodband, S (2016). Developing an inventory of core lexical bundles in English research articles: a cross-disciplinary corpus-based study. *Journal of World Languages*, 3(3), 184-203.
- Jiang, N. A., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(3), 433-445.
- 金澤佑 (編) (2020).『フォーミュラと外国語学習・教育:定型表現研究入門』くろしお出版.
- 小出凱渡 (2024).「英作文における学習者の定型表現使用の実態 —母語話者・学習者コーパスに基づく比較分析—」筑波大学大学院教育学学位プログラム修士論文 (未公開)
- 久保田竜子 (2018).『英語教育幻想』筑摩書房.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85-102.
- Mikajiri, N. (2022). Quantitative Analysis of Lexical Bundles Used in Spoken English and Japanese Elementary and Junior High School English Textbooks. *Journal of Corpus-based Lexicology Studies*, 5, 1-23.
- 文部科学省 (2017a).『小学校学習指導要領 解説 外国語活動・外国語編』Retrieved from https://www.mext.go.jp/content/20201029_mxt_kyoiku01-100002607_11.pdf
- 文部科学省 (2017b).『中学校学習指導要領解説 外国語編 英語編』Retrieved from https://www.mext.go.jp/content/20210531_mxt_kyoiku01-100002608_010.pdf
- 文部科学省 (2018).『高等学校学習指導要領解説 外国語編 英語編』Retrieved from https://www.mext.go.jp/content/1407073_09_1_2.pdf
- 中田達也 (2022).『英語は決まり文句が8割:今日から役立つ「定型表現」学習法』講談社.
- Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford University Press.
- Northbrook, J., & Conklin, K. (2018). "What are you talking about?" An analysis of lexical bundles in Japanese junior high school textbooks. *International Journal of Corpus Linguistics*, 23(3), 311-334.
- Northbrook, J., & Conklin, K. (2019). Is what you put in what get out? Textbook — derived lexical bundle processing in beginner English learners. *Applied Linguistics*, 40(5), 816-833.
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60-71.
- Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp.191-226). Longman.
- 阪上辰也 (2013).「日本語母語英語学習者のエッセイに見られる共起表現の分析」『広島外国語教育研究』第16号, 159-169.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt. (Ed.), *Formulaic Sequences:*

引用文献

- Acquisition, processing and use* (pp. 1–22). John Benjamins.
- Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2019). *Understanding formulaic language: A second language acquisition perspective*. New York: Routledge.
- Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, 12(3), 214–225.
- 鈴木陽子 (2022). 「モノローグ発話における日本人英語学習者の 法助動詞の使用」『明治学院大学教養教育センター紀要：カルチュラル』第16巻, 1–10.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford University Press.
- Yan, H. (2019). I think we should…: Investigating lexical bundle use in the speech of English learners across proficiency levels. *International Journal of Translation, Interpretation, and Applied Linguistics (IJTIAL)*, 1(2), 1–16.

資料1：SMにおける学習者の過剰使用語連鎖リスト(差分値の絶対値が大きい順).....

順位	項目	差分値	順位	項目	差分値
1	i agree with this	-4.1131	51	it is bad for	-0.0467
2	agree with this statement	-1.7835	52	i agree this idea	-0.0439
3	i agree with the	-1.3588	52	i i don t	-0.0439
4	with this statement because	-0.8821	52	so they have to	-0.0439
5	i disagree with this	-0.7835	55	don t don t	-0.0431
6	i think it is	-0.6074	56	so i agree this	-0.0399
7	agree with the statement	-0.5852	56	statement because i think	-0.0399
8	it is important for	-0.3964	56	t agree with the	-0.0399
9	i don t agree	-0.3953	56	this statement this is	-0.0399
10	is not good for	-0.3666	60	can learn a lot	-0.0395
11	i have two reasons	-0.3358	60	learn a lot of	-0.0395
12	disagree with this statement	-0.2907	60	we can we can	-0.0395
13	i don t like	-0.2529	63	have a right to	-0.0383
14	agree with this opinion	-0.2492	64	agree this opinion because	-0.0323
14	i agree this statement	-0.2492	64	agree with the opinion	-0.0323
16	don t agree with	-0.2108	64	because i don t	-0.0323
17	with the statement because	-0.1933	64	have two reasons first	-0.0323
18	it is important to	-0.1757	64	i think i think	-0.0323
19	with this statement i	-0.1442	64	some people don t	-0.0323
20	think it is important	-0.1340	64	with the idea that	-0.0323
21	this statement because i	-0.1294	71	so that s why	-0.0319
22	a lot of money	-0.1222	72	can learn how to	-0.0288
23	i don t i	-0.1214	73	agree this statement i	-0.0256
24	with this opinion because	-0.1154	73	i agree with that	-0.0256
25	people who don t	-0.1143	73	i i agree with	-0.0256
26	i think we should	-0.1138	73	need a lot of	-0.0256
27	agree with the idea	-0.1086	73	statement this is because	-0.0256
27	i agree this opinion	-0.1086	73	this statement i think	-0.0256
29	this statement i have	-0.1022	73	with this statement this	-0.0256
30	it is not good	-0.0839	80	it is good for	-0.0252
30	that s why i	-0.0839	81	i can t i	-0.0224
32	i don t want	-0.0827	82	have a lot of	-0.0210
33	agree with this idea	-0.0783	83	and second reason is	-0.0196
33	statement i have two	-0.0783	83	because if we have	-0.0196
35	t i don t	-0.0779	83	disagree with the idea	-0.0196
36	i disagree with the	-0.0752	83	first reason is that	-0.0196
37	agree this statement because	-0.0675	83	how to how to	-0.0196
37	with this idea because	-0.0675	83	i agree the opinion	-0.0196
39	it is it is	-0.0623	83	i agree with it	-0.0196
40	so i don t	-0.0576	83	i think we can	-0.0196
41	so i think we	-0.0575	83	if we have a	-0.0196
42	disagree with this opinion	-0.0571	83	opinion it is important	-0.0196
42	don t i don	-0.0571	83	people don t like	-0.0196
44	is very bad for	-0.0547	83	t agree with this	-0.0196
45	so we have to	-0.0527	83	think it is not	-0.0196
45	there are many people	-0.0527	83	this opinion because i	-0.0196
47	so i think it	-0.0514	83	we don t have	-0.0196
48	a lot of things	-0.0486	98	we can know how	-0.0192
49	is bad for our	-0.0483	99	that it is important	-0.0184
50	it is very important	-0.0480	100	is very important for	-0.0174

注．分析では少数点第9位までの差分値を求めているが，本資料では少数点第4位までの値を掲載。

資料2: SMにおける学習者の過少使用語連鎖リスト(差分値の絶対値が大きい順).....

順位	項目	差分値	順位	項目	差分値
1	i think it s	2.5798	53	i believe it s	0.0255
2	i don t think	1.3586	53	i don t believe	0.0255
3	to be able to	0.5752	53	think it would be	0.0255
4	don t think it	0.2825	56	they want to do	0.0254
5	think it s important	0.2676	57	believe that it s	0.0253
6	i don t know	0.1848	58	it depends on the	0.0234
7	t think it s	0.1434	58	um i think it	0.0234
8	i think that it	0.1283	60	i think they should	0.0214
9	and i think that	0.1189	60	it s important that	0.0214
10	at the same time	0.1168	60	should not be allowed	0.0214
11	think it s a	0.1150	60	the other hand i	0.0214
12	on the other hand	0.1104	64	will be able to	0.0212
13	while they are in	0.1058	65	have to worry about	0.0195
14	don t have to	0.1011	65	that it should be	0.0195
15	with the statement that	0.0941	65	they are going to	0.0195
16	don t think that	0.0896	68	and things like that	0.0191
17	i think that s	0.0885	68	to learn how to	0.0191
18	it s it s	0.0818	68	to take care of	0.0191
19	i believe that it	0.0807	71	and it s very	0.0177
20	they don t have	0.0804	71	disagree with the statement	0.0177
21	and i think it	0.0781	71	in the real world	0.0177
22	the people around them	0.0708	71	so i think that	0.0177
23	should be able to	0.0673	71	the people who are	0.0177
24	a lot of people	0.0631	71	they should be able	0.0177
25	it s very important	0.0622	71	when i was a	0.0177
26	if you don t	0.0621	78	i mean it s	0.0175
27	it s good for	0.0605	78	it helps them to	0.0175
28	think it s very	0.0572	80	don t want to	0.0164
29	you don t have	0.0566	81	as i was saying	0.0159
30	i do believe that	0.0481	81	s a good idea	0.0159
31	i also think that	0.0451	81	that it would be	0.0159
32	it s bad for	0.0446	81	think it s really	0.0159
33	that it s important	0.0398	85	be able to go	0.0143
34	in the united states	0.0396	85	don t believe that	0.0143
35	you know it s	0.0391	85	don t have a	0.0143
36	be able to have	0.0372	85	i agree that it	0.0143
36	i do think that	0.0372	85	i think it would	0.0143
36	i don t really	0.0372	85	it s really important	0.0143
36	it s a good	0.0372	85	no i don t	0.0143
36	think it should be	0.0372	85	t want to be	0.0143
36	think that it s	0.0372	85	the person who is	0.0143
42	but i don t	0.0368	85	you re going to	0.0143
43	not be able to	0.0345	95	and i don t	0.0128
43	they re going to	0.0345	96	because it helps them	0.0127
45	s very important for	0.0322	96	because they need to	0.0127
46	think it s good	0.0319	96	if you want to	0.0127
47	believe that it is	0.0299	96	people who do not	0.0127
47	it would be a	0.0299	96	should be allowed to	0.0127
49	in all public places	0.0297	96	that they have to	0.0127
50	a good idea for	0.0276	96	there should be a	0.0127
50	yes i think it	0.0276	96	what they want to	0.0127
52	it s going to	0.0273			

注. 96位が複数項目あったため、96位までの103項目を記載している。

注. 分析では少数点第9位までの差分値を求めているが、本資料では少数点第4位までの値を掲載。

資料3: SD における学習者の過剰使用語連鎖リスト(差分値の絶対値が大きい順)

順位	項目	差分値	順位	項目	差分値
1	i i don t	-0.8941	51	want to continue my	-0.0495
2	i don t like	-0.7983	52	uh uh i i	-0.0492
3	umm i don t	-0.5313	52	want to go to	-0.0492
4	i went to the	-0.4104	54	so it is very	-0.0467
5	so i want to	-0.3936	54	uh i can t	-0.0467
6	uh i don t	-0.3856	54	uh uh uh i	-0.0467
7	i don t have	-0.3659	57	don t i don	-0.0461
8	when i was a	-0.3641	58	a lot of things	-0.0440
9	i i want to	-0.3638	58	but i can t	-0.0440
10	uh i want to	-0.3546	58	so i want you	-0.0440
11	i want to i	-0.3260	58	uh yes uh i	-0.0440
12	i think it is	-0.3084	62	i don t go	-0.0412
13	i want to continue	-0.2781	63	uh i uh i	-0.0407
14	want to i want	-0.2581	64	i was a child	-0.0385
15	i can i can	-0.2411	65	yes yes yes yes	-0.0372
16	i want you to	-0.2081	66	when i when i	-0.0367
17	to i want to	-0.2034	67	it is it is	-0.0364
18	i went to your	-0.1961	68	but i want to	-0.0362
19	i i went to	-0.1811	69	are a lot of	-0.0355
20	don t want to	-0.1304	69	there are a lot	-0.0355
21	umm i want to	-0.1203	71	because i want to	-0.0340
22	i don t i	-0.1069	71	don t have enough	-0.0340
23	agree with this opinion	-0.1011	71	i want to teach	-0.0340
24	i think she is	-0.0983	74	don t like uh	-0.0336
25	i want to go	-0.0960	74	i want to study	-0.0336
26	uh i think uh	-0.0930	76	uh i i don	-0.0317
26	when i went to	-0.0930	76	want i want to	-0.0317
28	i don t want	-0.0920	78	if i if i	-0.0315
29	i want to do	-0.0893	79	and i want to	-0.0294
30	but i don t	-0.0855	79	i like japanese food	-0.0294
30	i agree with this	-0.0855	79	want you to return	-0.0294
32	with my family and	-0.0821	79	we can we can	-0.0294
33	i want to work	-0.0815	79	yes yes uh i	-0.0294
34	don t think so	-0.0783	84	i went to a	-0.0280
35	i agree with it	-0.0747	85	yeah yeah yeah yeah	-0.0275
36	i i can t	-0.0744	86	and it is very	-0.0273
36	i want to uh	-0.0744	87	i want to umm	-0.0272
38	so i have to	-0.0723	87	there are many people	-0.0272
39	t i don t	-0.0707	89	i like i like	-0.0254
40	i don t agree	-0.0679	90	it is not good	-0.0252
41	so i can t	-0.0614	90	to i have to	-0.0252
42	umm when i was	-0.0614	90	uh it s not	-0.0252
43	uh when i was	-0.0584	90	yes yes yes uh	-0.0252
44	i have to i	-0.0582	94	don t agree with	-0.0233
45	i can t i	-0.0553	94	i don t uh	-0.0233
45	i have i have	-0.0553	94	i i have to	-0.0233
47	i want i want	-0.0524	94	i think so too	-0.0233
48	i want to be	-0.0522	94	she is very strong	-0.0233
49	i have to earn	-0.0517	94	uh i agree with	-0.0233
50	uh i went to	-0.0495	100	i i didn t	-0.0231

注. 分析では少数点第9位までの差分値を求めているが, 本資料では少数点第4位までの値を掲載。

資料4: SDにおける学習者の過少語連鎖の差分値リスト(差分値の絶対値が大きい順).....

順位	項目	差分値	順位	項目	差分値
1	i think it s	1.9706	40	s going to be	0.0434
2	it s it s	0.8493	40	the same time i	0.0434
3	we had to leave	0.3794	40	well i think it	0.0434
4	i don t think	0.3651	40	yes i do it	0.0434
5	i am going to	0.2385	40	it s like a	0.0406
6	i would like to	0.2212	57	t think it s	0.0354
7	i don t really	0.1951	58	i didn t know	0.0348
7	yes i do i	0.1951	59	as you can see	0.0325
9	think it s a	0.1734	60	as you know i	0.0325
10	it s kind of	0.1707	60	don t know how	0.0325
11	so i think it	0.1539	60	i am trying to	0.0325
12	don t think it	0.1340	60	i ve been to	0.0325
13	at the same time	0.1328	60	it s a good	0.0325
14	no i don t	0.1153	60	not going to be	0.0325
15	but i think it	0.0976	60	or something like that	0.0325
15	i do it s	0.0976	60	they are going to	0.0325
15	it s a very	0.0976	60	would like to continue	0.0325
18	thank you very much	0.0917	60	so i think that	0.0275
19	don t have a	0.0830	70	they should be able	0.0271
20	it was it was	0.0813	71	and i don t	0.0251
21	my friend and i	0.0734	72	i have to pay	0.0246
22	don t know what	0.0677	73	a bit of a	0.0244
22	don t think i	0.0677	74	a it s a	0.0244
22	i am not a	0.0677	74	a little bit more	0.0244
22	think it would be	0.0677	74	and it s just	0.0244
22	to be able to	0.0677	74	and it s very	0.0244
27	a lot of the	0.0650	74	and it was it	0.0244
27	i d like to	0.0650	74	and they don t	0.0244
27	it s going to	0.0650	74	and we would like	0.0244
27	we would like to	0.0650	74	are going to be	0.0244
31	a lot of people	0.0616	74	because it s a	0.0244
32	should be able to	0.0569	74	but i am not	0.0244
32	uh huh uh huh	0.0569	74	but it s very	0.0244
34	because i don t	0.0546	74	can t afford to	0.0244
35	as long as they	0.0542	74	couldn t stand the	0.0244
35	i used to go	0.0542	74	didn t have any	0.0244
35	t know how to	0.0542	74	do it s a	0.0244
38	have a lot of	0.0483	74	don t i am	0.0244
38	they don t have	0.0483	74	don t think that	0.0244
40	am not going to	0.0434	74	had to leave the	0.0244
40	and i think i	0.0434	74	has a lot of	0.0244
40	going to be a	0.0434	74	i didn t see	0.0244
40	have yes i did	0.0434	74	i do have a	0.0244
40	i am more comfortable	0.0434	74	i do i have	0.0244
40	i am not going	0.0434	74	i go to a	0.0244
40	i have yes i	0.0434	74	i think you know	0.0244
40	i think it would	0.0434	74	i want to get	0.0244
40	i think that s	0.0434	74	i was able to	0.0244
40	i was working at	0.0434	74	i was in the	0.0244
40	i was working in	0.0434	74	i would i would	0.0244
40	it s a it	0.0434	74	i would prefer to	0.0244
40	s a it s	0.0434	74	if you have a	0.0244

資料4: SD における学習者の過少語連鎖の差分値リスト(差分値の絶対値が大きい順).....

順位	項目	差分値	順位	項目	差分値
74	in the first place	0.0244	74	think it s okay	0.0244
74	is going to be	0.0244	74	think it should be	0.0244
74	it s a really	0.0244	74	to finish our meals	0.0244
74	it s just a	0.0244	74	to pay for my	0.0244
74	it s just the	0.0244	74	uh let me see	0.0244
74	it was kind of	0.0244	74	was working in a	0.0244
74	it was really nice	0.0244	74	we weren t able	0.0244
74	know i don t	0.0244	74	well i i think	0.0244
74	like i don t	0.0244	74	well i think that	0.0244
74	maybe once a month	0.0244	74	weren t able to	0.0244
74	me and my friend	0.0244	74	yes i have yes	0.0244
74	no i didn t	0.0244	74	you can you can	0.0244
74	right now so i	0.0244	74	you don t have	0.0244
74	s it s a	0.0244	74	you know i am	0.0244
74	s kind of a	0.0244	74	you know i think	0.0244
74	so no i don	0.0244	74	you re going to	0.0244
74	so that s why	0.0244	74	you very much sir	0.0244
74	that s going to	0.0244	74	you want me to	0.0244
74	the number of people	0.0244			

注. 74位が複数項目あったため, 74位までの141項目を記載している。

注. 分析では少数点第9位までの差分値を求めているが, 本資料では少数点第4位までの値を掲載。

機械翻訳(MT)を学習ツールとするための考察

— 英作文を学ぶ大学生を対象とした量的・質的調査 —

研究者:湯浅 麻里子 大阪府／関西大学大学院 在籍

《研究助言者:寺内 一》

概要

本調査は、習熟度が低い大学生英語学習者がどのように機械翻訳(MT)を使っているのか、現状をマクロとミクロの視点で明らかにした。まず、773名の参加者による質問紙データを用いて、MTを使ってライティングを行う際のエンゲージメントを包括的に測る「MTエンゲージメント」尺度を作成した。18項目から成る尺度は、行動、プリエディット認知、ポストエディット認知、情意、社会の5つの因子で構成され、信頼性と妥当性が確認された。続いて尺度の5つの変数に基づき、434名の参加者を類似するグループに分けた。参加者は、①低エンゲージメント、②中エンゲージメント、③高エンゲージメントの3つのグループに分かれた。MTを使って英文を書くことに対する自己効力感、高エンゲージメント・グループが高く、中・低エンゲージメント・グループが低かった。さらにインタビューから、7名の学生がそれぞれの文脈の中で工夫しながらMTを使っていたことがわかった。

1 はじめに

機械翻訳(MT)は2016年、AIを用いたニューラル機械翻訳(NMT)の導入を機に、翻訳精度が劇的に向上した(Ducar & Schocket, 2018; Stapleton & Kin, 2019)。MTは私たちの日常に浸透し、グローバル社会の中で言語の壁を打ち破っている。その影響は教育現場も例外ではない。特に大学生の外国語(L2)ライティングにおけるMT使用は世界的に広がり(Briggs, 2018; O'Neill, 2019), “MT-enabled L2 writing” (MTを活用したL2ライティング; Jiang et al., 2024, p.1) が標準化しつつある。MTを取り入れることで、L2ライティングの質が高まり(Chon et al., 2021; Lee, 2020; Tsai, 2019), 語彙の使用と習得にも役立つことが報告されている(Lo, 2023)。これらの研究の多くは、ヨーロッパ言語共通参照枠(Common European Framework of Reference for Languages; CEFR)のB1以上の中・上級学習者を対象としており、MTは習熟度が高い学習者に効果があるとされている(Klimova et al., 2022)。一方、日本では大学生のおよそ8割がCEFR A2レベルで(Negishi et al., 2013), このような初級学習者を対象とした研究は世界でも限られている(Lee, 2021)。日本人大学生の多くが英語の課題にMTを使っているようだが(Yamada et al., 2021), 授業での活用例はまだ少ない。しかし、AI技術の急成長と実社会での役割を考えると、大学生のMT使用はもはや見過ごせない(Godwin-Jones, 2022; Kern, 2024)。

MTをL2学習に役立てるためには、入力時に「MTが出力しやすい母語(L1)」(“translation-friendly writing”)を考え、L1を修正するプリエディット(前編集)や、MTが出力したL2の文法や語彙の誤りを修正するポストエディット(後編集)などの方略を使うことが重要である(Bowker & Buitrago-Ciro, 2019)。つまり、ライティングのプロセスにどれだけ深く考えながらたずさわるか、「エンゲージメント」

が学習の鍵となる。しかし、日本の教育現場ではMT使用の事例も乏しく、学生のMTの使い方はあまり理解されていない(Gally, 2019)。そこで本調査は、エンゲージメントの枠組みを使い、大学生がどのようにMTを使っているのか、現状をマクロとミクロの視点から明らかにする。エンゲージメントとは、学習プロセスへ積極的に参加することで、行動、認知、情意、社会などの複数の要素で構成される(Hiver et al., 2021)。エンゲージメントはL2学習を促進することが報告されている(Lei et al., 2018; Liu et al., 2023)。大学生のMT使用の現状をエンゲージメントの観点から包括的に捉えることで、教師は学生のニーズや課題に気づき、教授法を見直すきっかけになるのではないかと。本調査では、非英語専攻の大学生の多くに相当するCEFR A2レベルの学習者を対象に、MTを使ってライティングを行うことへのエンゲージメント(「MTエンゲージメント」と略す)を測る尺度を作り、大学生の現状を数値で概観する。さらに個人に焦点を当ててインタビューを行い、学習者がそれぞれの文脈の中でどのようにMTを使っているのかを、より具体的に描写することを試みる。

2 先行研究

2.1 外国語学習とエンゲージメント

教育心理学から発展したエンゲージメント(Reschly & Christenson, 2022)は、近年、第二言語習得の分野で動機づけと学習行動を結びつける概念として注目を集めている(Hiver et al., 2021; Mercer & Dörnyei, 2020)。エンゲージメントは主に行動(努力や時間)、認知(深い思考や方略使用)、情意(学習に対する態度や感情)の3つの要素から構成され(Mercer, 2019)、学習者の行動、認知、情意がどのように促されて維持されるかを表す(Schunk & DiBenedetto, 2022)。さらに、社会(L2での他者とのやりとり; Svalberg, 2009)や主体性(学習への積極的な関わり; Reeve & Tseng, 2011)が構成要素に含められることもある。外国語を習得するには、語彙や文法を理解するだけでなく、持続的にL2を使用することが不可欠である(Hiver et al., 2024)。このように、エンゲージメントは学習行動を多面的な観点で捉えることができる。エンゲージメントとL2学習の成果には相関があることも実証されている(Lei et al., 2018; Liu et al., 2023)。

2.2 機械翻訳(MT)とエンゲージメント

習熟度が高い英語学習者がMT訳を参照しながら自作の英文を修正すると、語彙や文法の誤りが減り、高度な語彙や複雑な構文が増えることが報告されている(Chon et al., 2021; Lee, 2020; Tsai, 2019)。しかしMTは、L1を入力するだけで瞬時に英文ができる。MTを使うときは、ポストエディットを行うことが推奨されるが(Bowker & Buitrago-Ciro, 2019)、習熟度の低い学生はMT出力の誤りを見つけることが自体が困難である(Kol et al., 2018; Shin & Chon, 2023)。また、ポストエディット方略の使用頻度も習熟度に影響される(Chung, 2020; Shin & Chon, 2023)。日本人大学生のおよそ8割を占めるCEFR A2レベル(Negishi et al., 2013)の英語学習者によるMT使用は、学習へ悪影響を及ぼすと懸念されている。

コンピュータの登場以来、L2学習はテクノロジーの恩恵を受けてきた。私たちの日常に存在しているテクノロジーを排除することは学習環境を制限してしまう(Chun et al., 2016)。MTも例外ではない。L2学習にMTを取り入れることはごく自然ではないか。MTは学習者を惹きつけ、エンゲージメントを高めるタスクを作ることができる(Kern, 2024)。学習者は有意義で価値あるタスクには意欲的に取り組む(Williams & Burden, 1997)。このような学習機会を作るためには、教師は大学生がどのようにMTを使ってライティングをしているのか、現状を知ることが不可欠である。エンゲージメントの枠組みを使うことで、学生のMT使用の実態を行動、認知、情意、社会などのさまざまな観点から包括的に理解することができるだろう。

2.3 エンゲージメントと自己効力感

学習に大きな影響を与える要因の1つである自己効力感とは、エンゲージメントと強く関係している。自己効力感とは、ある課題に対する自身の能力へのビリーフで、自分はできると信じることである(Bandura, 1997)。このビリーフが課題の選択や努力に影響を与え、学習成果を高める(Pajares, 1996)。教育心理学では、自己効力感と動機づけ、エンゲージメント、学習成果との強い相関が実証されている(Schunk & DiBenedetto, 2022; Zimmerman, 1995)。

教育心理学における自己効力感の知見は、第二言語習得においても広く研究されており、自己効力感と外国語4技能の習得には強い相関があることがわかっている(Goetze & Driver, 2022; Wang & Sun, 2020)。自己効力感は単独で機能するのではなく(Graham, 2022)、自己調整学習や学習方略などと連動して機能する(Bai & Wang, 2023; Sun & Wang, 2020)。リーディングやライティングなどの方略指導が自己効力感を高めることも実証されている(Graham et al., 2020; Teng & Zhang, 2020)。エンゲージメントと自己効力感においても正の相関がある。Cai and Xing(2023)は、L2学習に対する自己効力感が、認知エンゲージメントを通じて成果を高めることを示している。Tsao(2021)は、L2ライティングに対する自己効力感が、フィードバックに対するエンゲージメントを高めることを報告している。自己効力感がエンゲージメントを高め、パフォーマンスが向上し、達成感や自己効力感を促すことで、さらなる学習の好循環が生み出されるのである。

2.4 調査の目的と構成

本研究は、CEFR A2レベルに相当する非英語専攻の大学生英語学習者が、MTを使ってどのようにL2ライティングを行っているのかを、マクロとミクロの視点から調査する。調査1では、CEFR A2レベルのL2学習者に特化した「MTエンゲージメント」尺度を作成する。調査2では、大学生のMT使用の実態を調査1で作成した尺度による量的手法と、インタビューによる質的手法により明らかにする。

3 調査1: MTエンゲージメント尺度の作成

3.1 目的

調査1では「MTエンゲージメント」尺度を作成する。エンゲージメントは個々の学習者が置かれた文脈に依存する(Hiver et al., 2024)。妥当性の高い尺度を開発するため、2つのリサーチクエスション(RQ)を設定した。

RQ1 MTエンゲージメント尺度はどのような因子で構成されるか。

RQ2 MTエンゲージメント尺度の信頼性と妥当性はどの程度か。

3.2 方法

3.2.1 参加者

調査者が所属する大学の17学部773名の学生(男性387名、女性372名、回答しない14名)が本調査への協力を同意し、参加した。参加者は非英語専攻の学生を対象とした必修英語クラスの1年生(392名)と2年生(381名)で、英語習熟度の判定には、GTEC-LRによるプレスメントテスト(500点満点)の結果($N = 717$,

$M = 220.65$, $SD = 46.91$)を利用した。点数の分布からおおよそ9割がCEFR A2レベルを占め、少数のA1とB1レベルを含んでいた。参加者は平均的な非英語専攻の大学生英語学習者を表していると言える。

データ収集に際しては、回答は無記名で、データはすべて統計的に処理し、個人が特定されないことを説明した。調査への協力に同意した学生はスマートフォンを使ってGoogle Forms から回答した。データ収集に先立ち、調査者が所属する大学院の倫理委員会に許可を得た。

3.2.2 測定道具

本調査では2つの質問紙を用いた。本調査のために作成した「MTエンゲージメント」尺度と「MTを用いたL2ライティングに対する自己効力感(「MT自己効力感」と略す)」尺度(Yuasa & Takeuchi, 2022)である。エンゲージメントが自己効力感とどの程度の相関があるか(併存的妥当性)を調べるためである。

3.2.3 実施手順

MTエンゲージメント尺度の作成は、項目の作成、尺度の作成、尺度の検証の3段階で行った(Boateng et al., 2018)。尺度の作成に探索的因子分析を、尺度の検証に検証的因子分析を実施した。第1段階では、MTエンゲージメントの5つの下位概念を定義した。「MTを使う時間と努力」を表す行動、「MT出力を向上させる方略使用」を表す認知、「MT使用への態度や感情」を表す情意、「学習者間のMTについてのやりとり」を表す社会、「積極的なMT使用と個人化」を表す主体性である。これらの定義と質的データによる学習者の方略使用に基づき(Yuasa & Takeuchi, in press), 30項目の草案を作成した。これらの項目は、第二言語習得分野の研究者と議論を通して修正を重ね、内容的妥当性を高めた。

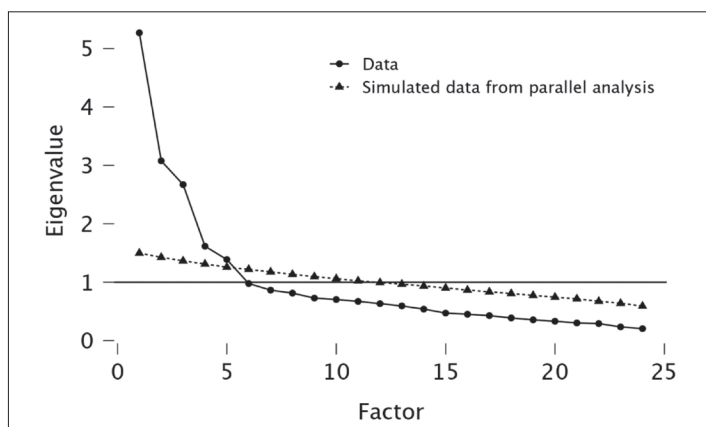
第2段階は尺度の作成である。2023年6月に108名、7月に269名の学生を対象に予備調査を実施した。その結果を受けて項目の改善を行い、24項目の質問紙を作成した。この質問紙を同年9月から10月に773名の学生に実施した。「あなたは普段、英文を書くときにどのように機械翻訳を使っていますか」という教示文に対して5件法(1. 当てはまらない 2. あまり当てはまらない 3. どちらとも言えない 4. やや当てはまる 5. 当てはまる)での回答を依頼した。データクリーニングにより、テストスコアの欠如や不自然な回答(すべての項目が同じ回答)、外れ値などを除いた708名のサンプルを、探索的因子分析と検証的因子分析のために、354名ずつの2グループに分けた。各グループ100名ずつを無作為抽出して t 検定を行い、習熟度に差がないことを確認した($t = .55$, $df = 198$, $p = 0.59$)。354名のデータを基に、因子分析の前提条件である、サンプル数が十分であること(KMOの標本妥当性指標 = .82)と、変数間に相関があること(バートレットの球面性検定: $\chi^2 = 3238.71$, $df = 276$, $p < .001$)を確かめた後、JASP(ver.0.18)を用いて探索的因子分析を実施した。因子負荷量の最低値はHair et al.(2019)に基づき.50とした。

第3段階では、探索的因子分析で特定した因子構造の内的―貫性と尺度の妥当性を、検証的因子分析により検証した。尺度項目の内的―貫性にはクロンバック α の信頼性係数を、また、妥当性は構成概念妥当性、収束的妥当性、弁別妥当性を調べた。さらにMT自己効力感尺度を用いて、外的な基準との間にどの程度相関があるのかを検証する併存的妥当性を調べた。

3.3 結果

3.3.1 MTエンゲージメントの因子構造(RQ1)

記述統計を基に、MTエンゲージメント質問紙のすべての項目の歪度と尖度が $\pm < 1.96$ (Hair et al., 2019)であることから、データが正規分布していることを確かめた。続いて探索的因子分析(最尤法・オリミン回転)を実施して因子を抽出した。固有値1以上、スクリープロット、平行分析の結果(図1)から、5因子構造を仮定した。表1に各質問項目と因子負荷量を示す。



■図1: MTエンゲージメント因子抽出のスクリープロット

■表1: 質問項目と探索的因子分析の結果

No.	項目	因子					独自性
		1	2	3	4	5	
4	英文を書くことについては、私は機械翻訳のヘビーユーザーである	.87	.05	-.04	.01	.00	.22
2	英文を書くときは機械翻訳を使うことが多い	.85	-.02	.06	.02	-.01	.27
5	機械翻訳は英文を書くのに欠かせない	.78	-.02	-.06	.01	.01	.37
3	自分で英文をうまく書けないと思ったときは、機械翻訳を積極的に使う	.64	.06	.10	.06	-.02	.51
1	和英辞書の代わりに機械翻訳を積極的に使っている	.55	-.09	.04	.09	.00	.68
17	機械翻訳を使うときは、意図した英文が出力されるまで日本語を修正する	.09	.71	-.05	-.03	-.01	.51
15	機械翻訳を使うときは、わかりやすい日本語で入力する	-.02	.63	.05	-.05	-.08	.60
16	機械翻訳を使うときは、日本語の主語を補いながら入力する	.05	.63	.01	-.03	-.09	.60
19	機械翻訳の使い方について、自分なりに試行錯誤してみる	-.04	.54	.10	.20	.13	.53
14	機械翻訳を使うときは、何通りか日本語を入力してみる	.12	.41	.25	-.06	.15	.61
8	機械翻訳で出力された英文に誤りがないかチェックし、必要があれば英語を直す	-.03	-.00	.82	-.02	-.02	.33
6	機械翻訳で出力された英文が正しいかどうか、読み返す	.01	-.01	.78	.02	-.02	.39
7	機械翻訳に入力した日本語と、出力された英語を必ず見比べる	.05	.01	.77	-.01	-.00	.38
22	機械翻訳を使うと英文を書くのが苦痛でなくなる	-.00	-.05	.01	.93	-.01	.16
21	機械翻訳を使うとスラスラ英文が書けるので気分がいい	.16	.07	-.11	.71	-.01	.40
23	機械翻訳を使えるなら、喜んで英文を書く	-.07	.05	.11	.55	.09	.65
11	自分が使っている機械翻訳以外に類似のアプリがあるか、友人や教員、学外の人に聞いてみる	.03	.02	-.01	-.02	.74	.46
10	機械翻訳の使い方について、友人や教員、学外の人に聞いてみる	-.01	.01	-.02	-.03	.70	.52
13	機械翻訳の使い方について、友人や教員と議論をする	-.06	-.03	-.04	.03	.68	.54
12	友人と一緒に機械翻訳を使って英文を書く	.34	-.13	-.02	-.02	.46	.68
9	機械翻訳で出力された英文の中に知らない単語があれば、その意味や用法について、辞書で調べる	-.09	.14	.34	.00	-.01	.83
18	機械翻訳の画面上にあるさまざまな機能を使ってみる	-.06	.36	.07	.20	.25	.67
20	異なる機械翻訳を試して、自分に合うものを探す	-.04	.32	.02	.21	.25	.71
24	機械翻訳の使い方について知りたい	-.15	.13	.03	.35	.20	.77
累積因子寄与率		.13	.22	.31	.40	.49	

注: 因子負荷量は > .40, 太字は因子負荷量 > .50 を表す。

分析の結果から、十分な因子負荷量(>.50)を示さなかった6つの項目(9, 12, 14, 18, 20, 24)は除外した。続いて因子を名づけた。第1因子はMTを積極的かつ頻繁に使う項目が多いことから「行動」とした。方略使用に関する項目は、第2因子の、入力時にMTが訳しやすいように日本語を修正する項目と、第3因子の、MT出力の誤りを修正する項目に分かれたため、前者を「プリエディット認知」、後者を「ポストエディット認知」と名づけた。第4因子はMT使用への態度と感情を表しているため「情意」とした。第5因子はMTについてのやりとりを表しているため「社会」とした。これらの因子は先行研究のエンゲージメントの理論と一致したが、主体性を想定した項目18と20は十分な因子負荷量を得られず、項目19は認知の前方略に分類されたため、主体性に関する因子は抽出されなかった。

3.3.2 MTエンゲージメントの信頼性と妥当性(RQ2)

RQ1で抽出された因子構造の信頼性と妥当性を調べるため、最尤法による検証的因子分析を行った。表2にMTエンゲージメント尺度の因子構造(因子負荷量、残差、信頼性係数、AVE、因子間相関)を示す。項目1, 13, 17, 19, 23は残差が.40より大きい(Hair et al., 2019)、因子負荷量が.50を超えているため残した。これらは項目1の「和英辞書の代わりにMTを積極的に使っている」など、質的データから得られた重要な特徴が含まれているためである(Yuasa & Takeuchi, in press)。その結果、1つの因子を構成するために最小限必要な3つの項目を保持した(Hair et al., 2019)。内的一貫性を検討するクロンバックの α 係数は、5つの因子すべてで.73から.89と閾値である.70を満たした。次にIn'nami and Koizumi(2011)を基にモデルの適合基準を用いて構成的妥当性を検証した。 $\chi^2 = 265.34(df = 125, p < .01)$, RMSEA = .06(90%信頼区間:.05–.07)(< .08), CFI = .95(> .90), TLI = .94(> .90), SRMR = .05(< .08)となり、良好な適合指数(Brown, 2015)が得られた。収束妥当性を示すAVEはプリエディット認知が.41と若干低いが、4

■表2: 検証的因子分析によるMTエンゲージメント尺度の因子構造

因子	項目番号	因子負荷量	残差	α 係数	AVE	因子間相関			
						1	2	3	4
1. 行動	4	.88	.23	.89	.64	-			
	2	.85	.27						
	5	.84	.29						
	3	.77	.41						
	1	.60	.65						
2. プリエディット認知	16	.79	.38	.73	.41	.44***	-		
	15	.76	.43						
	17	.53	.73						
	19	.51	.74						
3. ポストエディット認知	6	.84	.30	.86	.68	.12	.57***	-	
	7	.82	.32						
	8	.81	.34						
4. 情意	22	.81	.34	.75	.51	.42***	.23***	.06	-
	21	.78	.39						
	23	.56	.69						
5. 社会	11	.80	.37	.77	.55	.01	.09	.09	.25***
	10	.76	.43						
	13	.63	.61						

注:*** = $p < .001$. AVE = average variance extracted.

つの因子は期待値である .50を満たした(Hair et al., 2019)。さらに各因子のAVEの平方根を計算したところ、すべての値が因子間相関よりも大きかったことから弁別妥当性が確認された。最後に、併存的妥当性を評価するため、MTエンゲージメント尺度とMT自己効力感尺度($\alpha = .80$)の相関を調べたところ、ピアソン相関係数は $r = .36$ ($p < .001$)と中程度に近く、併存的妥当性は弱く支持された。したがって、MTエンゲージメント尺度の信頼性と妥当性は共に満たされていることが確かめられた。

4 調査2: MTエンゲージメントの量的・質的調査

4.1 目的

調査2では大学生英語学習者がMTを使って英文を書くときのエンゲージメントをマクロとミクロの視点から調査する。説明的順次デザインによる混合法(“explanatory sequential mixed methods design”; Creswell & Creswell, 2023, p. 16)を採用し、質問紙による量的データの分析後に、インタビューによる質的データを加えて解釈をする。数値では見えない個人に焦点をあてることで(Mackey & Gass, 2022)、全体像をより鮮明に記述するためである。量的データの分析にはクラスタ分析を用いる。クラスタ分析とは、あるサンプルを変数の組み合わせが似たグループ(クラスタ)に分類する多変量解析法である(Dörnyei, 2007)。第二言語習得研究では、動機づけや学習方略などの個人差要因によって学習者をグループに分けることで、学習者がたどる異なる学習経路を表すために利用されている(Staples & Biber, 2015)。調査1で作成した5因子のMTエンゲージメント尺度の回答データをクラスタ分析し、各因子の強さにより学習者を分類することで、グループ毎の学習者の特徴を明らかにする。さらにグループから数名ずつインタビューを行うことで、MTエンゲージメントをより具体的に記述する。そこで以下のリサーチクエスチョン(RQ)を設定した。

RQ1

大学生英語学習者は MTエンゲージメントの5因子(行動、プリエディット認知、ポストエディット認知、情意、社会)の強さによって、どのようなグループに分かれるか。

RQ2

これらのグループの MT自己効力感はどのように異なるか。

RQ3

各グループの学生はどのようにMTを使っているか。

4.2 方法

4.2.1 参加者

参加者は、調査への協力に同意した、調査者が所属する大学の15学部448名の学生(男性234名、女性214名)である。非英語専攻の学生を対象とした必修英語クラスの1年生(192名)と2年生(256名)で、英語習熟度の判定にはGTEC-LRによるプレメントテスト(500点満点)を利用した。テストの結果($N = 434$, $M = 190.30$, $SD = 52.26$)、点数の分布からおおよそ9割がCEFR A2レベルを占め、少数のCEFR A1とB1レベルを含む、平均的な非英語専攻の大学生英語学習者を表していた。

データ収集に際しては、質問紙への回答は無記名とし、データはすべて統計的に処理して個人が特定されないことを説明した。調査への協力に同意した学生はスマートフォンを使ってGoogle Formsに回答した。さらに回答した参加者からインタビューへの協力者を募集したところ、7名の学生が応じた。インタビュー参加者には実施の当日、調査の目的、データの収集方法、個人情報保護や匿名性などを紙面と口頭で説明し、書面による同意を得た。データ収集に先立ち、調査者が所属する大学院の倫理委員会

に許可を得た。

4.2.2 実施手順

調査2は、質問紙による量的調査とインタビューによる質的調査の2つの調査で構成される。量的調査では、2023年11月から12月にかけて、調査1で作成した「MTエンゲージメント」尺度と「MT自己効力感」尺度の2つの質問紙を配布して回答データを得た。質問紙の尺度は分散を広げて差を可視化するため、6段階尺度(1. まったく当てはまらない 2. 当てはまらない 3. あまり当てはまらない 4. やや当てはまる 5. 当てはまる 6. とてもよく当てはまる)を採用した。学年、学部、性別などの基本情報のほかに、MTの使用頻度を6段階(1. まったく使わない 2. ほとんど使わない 3. あまり使わない 4. ときどき使う 5. よく使う 6. いつも使う)で尋ねる項目も設けた。質的調査では2024年1月から2月にかけて、7名の協力者に半構造化インタビューを対面およびZoomで実施した。所要時間は1人およそ30分を予定し、合計180.9分(平均25.8分)の音声データを得た。

4.2.3 分析方法

質問紙回答から得たデータはJASP(ver.0.18)を使い、階層クラスタ分析(ユーグリッド距離・ワード法)により回答者をエンゲージメントの強さによりグループ分けした(RQ1)。その後、分散分析によりグループ間に差があることを確かめた。さらに各グループの自己効力感に違いがあるかを調べるために、分散分析を行った(RQ2)。続いて、学生個人のMT使用を調べる(RQ3)ために、半構造化インタビューで得た音声データを自動文字起こしツール(Notta AI)と手作業を併用してテキスト化した。テキストデータからMTエンゲージメントを表す部分を抽出し、内容に基づいて6つのエンゲージメントコード(行動、プリエディット認知、ポストエディット認知、情意、社会、主体性)を作成し、定義した。定義に基づき、7人から抽出した合計112個のテキストデータの抜粋に6つのコードでラベルづけした。客観的な判断を得るために、第二言語習得分野に属さないもう1人の分析者がラベルづけを行い、2者間の分析の一致度を確かめた。その結果、106個(94.6%)のラベルが一致し、6つの相違項目は話し合いによって合意に至った。

4.3 結果

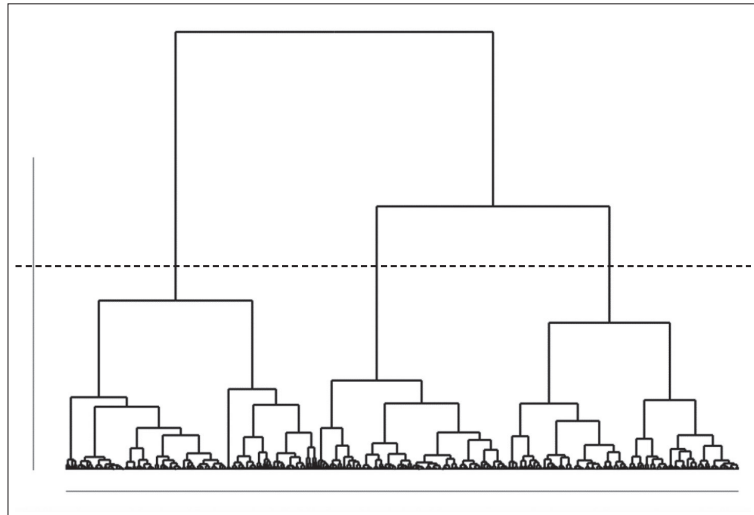
4.3.1 MTエンゲージメントによるグループ分け(RQ1)

データクリーニングにより、テストスコアの欠如や外れ値などを除いた434名のサンプルを使用した。MTの使用頻度を6段階でたずね、その結果、いつも使う($n = 63$)、よく使う($n = 215$)、ときどき使う($n = 132$)の合計が410名で全体の94.5%を占めた。一方、あまり使わない($n = 12$)、ほとんど使わない($n = 8$)、まったく使わない($n = 4$)の合計は24名(5.5%)だった。程度の差はあるが、調査対象となった学生の94.5%がMTを使っているようだ。まったく使わないと回答した学生も、質問紙の回答に1の「まったく当てはまらない」以外の選択肢を含んでいたため、すべてのデータを分析に用いた。表3に質問紙回答から得た各項目の平均と標準偏差を示す。行動とプリエディット・ポストエディットの両認知エンゲージメントは平均値が高かった。情意エンゲージメントは平均値が比較的高い傾向が見られたが、社会エンゲージメントは他の変数に比べると際立って低かった。

■表3: MTエンゲージメント尺度の項目と平均・標準偏差

No.	質問紙項目	M	SD
行動エンゲージメント($\alpha = .89$)			
4	英文を書くことについては、私は機械翻訳のヘビーユーザーである	4.37	1.18
2	英文を書くときは機械翻訳を使うことが多い	4.53	1.03
5	機械翻訳は英文を書くのに欠かせない	4.50	1.10
3	自分で英文をうまく書けないと思ったときは、機械翻訳を積極的に使う	4.85	0.96
1	和英辞書の代わりに機械翻訳を積極的に使っている	4.76	1.10
プリエディット認知エンゲージメント($\alpha = .70$)			
17	機械翻訳を使うときは、意図した英文が出力されるまで日本語を修正する	4.26	1.18
15	機械翻訳を使うときは、わかりやすい日本語で入力する	4.75	1.07
16	機械翻訳を使うときは、日本語の主語を補いながら入力する	4.68	1.06
19	機械翻訳の使い方について、自分なりに試行錯誤してみる	3.99	1.08
ポストエディット認知エンゲージメント($\alpha = .84$)			
8	機械翻訳で出力された英文に誤りがないかチェックし、必要があれば英語を直す	4.77	1.05
6	機械翻訳で出力された英文が正しいかどうか、読み返す	4.84	1.06
7	機械翻訳に入力した日本語と、出力された英語を必ず見比べる	4.80	1.02
情意エンゲージメント($\alpha = .83$)			
22	機械翻訳を使うと英文を書くのが苦痛でなくなる	4.01	1.14
21	機械翻訳を使うとスラスラ英文が書けるので気分がいい	3.85	1.16
23	機械翻訳を使えるなら、喜んで英文を書く	3.61	1.33
社会エンゲージメント($\alpha = .81$)			
11	自分が使っている機械翻訳以外に類似のアプリがあるか、友人や教員、学外の人に聞いてみる	2.86	1.32
10	機械翻訳の使い方について、友人や教員、学外の人に聞いてみる	2.63	1.26
13	機械翻訳の使い方について、友人や教員と議論をする	2.52	1.32

続いて5つのMTエンゲージメントの強さによって参加者をグループに分けるため、階層クラスタ分析(ユークリッド距離・ワード法)を行った。デンドログラム(図2)と各グループの人数、解釈が可能かどうかの判断から、3クラスタによる分類を採用した。



■図2: MTエンゲージメント・クラスタのデンドログラム

3つのグループ間の差を調べるための分散分析を行った。表4にグループ毎の各変数の平均と標準偏差, F 値, 効果量を示す。

■表4: 分散分析の結果

変数	グループ1 ($n=136$) $M(SD)$	グループ2 ($n=100$) $M(SD)$	グループ3 ($n=198$) $M(SD)$	F 値	効果量 ¹ (η^2)	チューキーの 多重比較
行動	4.05 (1.00)	4.35 (0.69)	5.10 (0.59)	83.76***	.28	1-3, 2-3
プリエディット認知	3.80 (0.71)	4.42 (0.65)	4.85 (0.61)	103.26***	.32	1-2, 1-3, 2-3
ポストエディット認知	3.95 (0.85)	5.03 (0.64)	5.28 (0.60)	154.12***	.42	1-2, 1-3
情意	3.55 (0.91)	2.96(0.77)	4.47 (0.82)	118.40***	.36	1-2, 1-3, 2-3
社会	3.08 (0.96)	2.01(0.75)	2.72 (1.20)	31.16***	.13	1-2, 2-3

注: *** = $p < .001$ 1効果量: $> .14$ = 大, $.06 \sim .14$ = 中, $< .06$ = 小。

3つのエンゲージメント・グループは, すべての変数に有意差が見られた。効果量については, 行動, 認知(プリエディット・ポストエディット), 情意が大に対して, 社会のみが中程度だった。これらの結果に基づき, 3つのグループの特徴を解釈し, グループ1を「低エンゲージメント」, グループ2を「中エンゲージメント」, グループ3を「高エンゲージメント」と名づけた。表5にグループの特徴を示す。

■表5: MTエンゲージメント・グループの特徴

グループ名	人数 (割合)	MT使用頻度	特徴
1. 低エンゲージメント	$n=136$ (31.3%)	$M = 4.29$ ($SD = 1.01$)	社会以外のエンゲージメントが低い。 特に行動と認知がグループの中で一番低い。
2. 中エンゲージメント	$n=100$ (23.0%)	$M = 4.58$ ($SD = 0.70$)	行動と認知は3グループの中間だが, 情意と 社会がグループの中で一番低い。
3. 高エンゲージメント	$n=198$ (45.6%)	$M = 5.03$ ($SD = 0.74$)	社会以外のすべてのエンゲージメントが グループの中で一番高い。

グループの特徴は、高エンゲージメント・グループがサンプルの半数近くを占め、4つのエンゲージメント変数とMTの使用頻度が一番高かった。このグループは積極的にMTを使っている可能性がある。一方、全体の3分の1弱を占める低エンゲージメント・グループは、行動と認知が一番低く、MT使用頻度も相対的に低いことから、MT使用にはそれほど積極的でない可能性がある。中エンゲージメント・グループは情意と社会が一番低いことから、MTを使っているが、MTに対してあまり前向きな態度ではない可能性がある。

4.3.2 MTエンゲージメントによるグループの自己効力感(RQ2)

3つのMTエンゲージメント・グループの自己効力感に対する差を調べた。表6にMT自己効力感尺度の質問紙項目と、回答から得た各項目の平均と標準偏差を示す。

■表6: MT自己効力感尺度の項目と記述統計

No.	質問紙項目 ($\alpha=.84$)	M	SD
1	機械翻訳を使って英文を書くとき、正しい文法で書くことができると思う	4.01	0.98
2	機械翻訳を使って英文を書くとき、自然な文章を書くことができると思う	3.54	1.05
3	機械翻訳を使って英文を書くとき、読み手に簡潔でわかりやすい英文を書くことができると思う	3.94	1.03
4	機械翻訳を使って英文を書くとき、自分の意見や考えを適切に表現することができると思う	3.79	1.08
5	機械翻訳を使って英文を書くとき、適切な単語を選ぶことができると思う	3.67	1.07

グループ毎の平均は、低エンゲージメント ($M=3.65$, $SD = 0.74$), 中エンゲージメント ($M=3.34$, $SD = 0.74$), 高エンゲージメント ($M=4.12$, $SD = 0.76$) となった。1元配置の分散分析を行った結果、3グループ間の平均の差は有意となった [$F(2, 431) = 39.25$ ($p < .001$, $\eta^2 = .15$)]。チューキー法による多重比較を行ったところ、低エンゲージメントと高エンゲージメント、中エンゲージメントと高エンゲージメントの平均値に有意な違いが見られた。高エンゲージメント・グループのMT自己効力感が一番高く、中・低エンゲージメント・グループは低かった。

4.3.3 インタビューによる学生個人のMT使用(RQ3)

MTエンゲージメントの各グループ内の学生がどのようにMTを使っているのかを、半構造化インタビューにより明らかにした。インタビューでは事前に準備した質問内容(資料1)を引き出すことを念頭におきながら、調査者と初対面の学生がリラックスして話せるよう、自然な対話の流れを心がけた。協力者の7名の学生は、中エンゲージメント・グループに3名、高エンゲージメント・グループに4名が属した。低エンゲージメント・グループに属するインタビュー参加者はいなかった。表7に参加者のプロフィールを示す。

■表7: インタビュー参加者のプロフィール

グループ	名前*	性別	年齢	学年	GTEC L-R**	よく使うMTツール
中エンゲージメント	Aoi	女	19	2	201	Google 翻訳, DeepL
	Mika	女	19	2	226	Google 翻訳
	Saeko	女	20	2	195	Google 翻訳
高エンゲージメント	Koki	男	19	1	79	Google 翻訳
	Naomi	女	19	1	105	Google 翻訳
	Taka	男	19	1	89	Google 翻訳
	Yoko	女	19	1	119	Google 翻訳, Patago

*仮名を使用。 **GTEC L-R スコアは500点満点

参加者は全員、Google翻訳を使っていたが、Aoi と Yoko はそれぞれ別の MT ツール (DeepL と Patago) を併用していた。インタビューからより詳細なエンゲージメントが明らかになった。表8～表13 にテキストデータの抜粋を要約した。名前の横の数字は会話のやりとりの順番を示す。

行動エンゲージメント (どんなときに、もしくは何のために MT を使うか) では、英語のライティング課題だけでなく、さまざまな理由で MT を使っていた (表8)。

■表8: 行動エンゲージメント

中エンゲージメント	高エンゲージメント
・寮の留学生と交流するとき (Aoi, 54) ・バイト先で外国人客の対応に、覚えているフレーズで対処できないとき (157)。仏語の教科書から使える動詞・名詞を探すため (Mika, 161) ・オンライン授業について行くため (Saeko, 100)	・ラグビーの国際試合で海外の選手と交流するとき (Koki, 75) ・英語日記を書くときにより良い表現を探すため (Naomi, 198) ・英作の課題をするとき (Taka, 116) ・英作文の課題でうまく表現できないとき (Yoko, 110)

Aoi, Mika, Koki は外国人とコミュニケーションを図るために MT を使っていた。Aoi は寮での留学生との会話に MT を使っていた。以下にインタビューの抜粋を示す。() 内は調査者が補足した。

抜粋 1

機械翻訳を使わずに自分の英語力で話してみようって思うこともあるんですけど、会話が途切れたり…相手も「え？」って戸惑っているときは、もうすぐに機械翻訳を使っています (Aoi, 140)。

Saeko は1年次に、コロナ禍以降も続いたネイティブ講師によるオンライン授業について行くため、必死に MT を使っていたことを語った (抜粋2)。

抜粋 2

(大学に)入学したとたん、(授業で)英語で書いたり話したりしようって。コミュニケーションの練習をしたことなかったのに。もう(パソコンの)横で翻訳しながらしゃべらないと追いつけないし (100)。そこからどんどん使うようになって便利さもわかってしまって (Saeko, 102)。

7人の中で Mika, Naomi, Yoko は自分で書いた英文を改善するために MT を使っていた。英語で「うまく表現できないときのモヤモヤ感」(27)が嫌いな Mika は、理由をこう説明した (抜粋3)。

抜粋 3

自分の中にない表現を使うときに、やっぱり調べないと(英語の)知識は増えないから (79)。受験のとき以外は、多分、何を学ぶかより、どうやって調べるかのほうが大事だと思います (Mika, 174)。

プリエディット認知エンゲージメントからは、全員が何通りかの日本語を入力し直していることがわかった (表9)。

■表9: プリエディット認知エンゲージメント

中エンゲージメント	高エンゲージメント
・書き言葉で入力する (Aoi, 88) ・日本語の表現を色々変えてみる (Mika, 109) ・わかりやすい日本語で入力する (151)。使いたい構文を考えて日本語を入れる (Saeko, 134)	・意図した英語が出るまで日本語を入力する (Koki, 133) ・教科書のような日本語で入力する (Naomi, 167) ・日本語の語順(副詞の位置など)を変える (Taka, 142) ・シンプルで形式的な日本語で入力する (Yoko, 137)

特に Aoi, Mika, Saeko, Naomi, Yoko の5名は、簡潔でわかりやすい日本語で入力する、口語ではなく書き言葉で入力するなど、MT が出力しやすい日本語 (translation-friendly writing) を意識していた。

ポストエディット認知エンゲージメントから、参加者全員がMTの英文を読んで確認していることがわかった(表10)。

■表10: ポストエディット認知エンゲージメント

中エンゲージメント	高エンゲージメント
- 日本語を Google 翻訳に入力し、その出力を DeepL に入れ、日本語が適切か確かめる (Aoi, 72) - 知らない語は知っている語に置き換えたり (97)、ウェブで調べる (105)。英語の受動・能動態を変える (101)。逆翻訳で意図した意味か確かめる (Mika, 123) - 意図した構文が出ない、英文が長すぎる (134)。ニュアンスが違うとわかったら、日本語を修正する (Saeko, 137, 153)	- 合っているかわからないけど、時々時制を変える (Koki, 202) - 逆翻訳でニュアンスを確かめる (115)。単語をウェブで調べたり置き換える (121)。文法を直す (Naomi, 161) - いつも一応、自分で英文を確認してから使う (Taka, 121) - Patago の訳出を Google 翻訳に入れ、日本語が合っているか確かめる (86)。知らない英単語は日本語に訳したり (115) ウェブで調べる (Yoko, 125)

参加者は全員、MTの出力は不完全だという認識を持っており、意図した英文になっているか、逆翻訳機能を使ってMTの出力を日本語に戻して内容を確認していた。AoiとYoko はさらに別のMTツールを使って意味が通るかどうかを確認していた。Aoi は寮でのカジュアルな会話には Google 翻訳の方が良いと言いつつ、DeepL を併用している理由を説明した(抜粋4)。

抜粋4

逆翻訳も使いますが、Google だけだったら、DeepL で確かめたときにちょっと言い方がきつくなっていたことがあって(74)。… 英語でも丁寧な言い方とかあると思うので、まず丁寧な日本語で入れるようにして、ほんとうにそれで丁寧な言い方になっているのかは、やっぱり (Google から DeepL に) コピペで確かめます (Aoi, 88)。

文法の修正については、時制を変える (Koki) から、受動態・能動態など文の構造を変える (Mika) まで、ポストエディットには程度に差が見られた。インタビューでは Saeko と Naomi が方略の説明に「ニュアンス」という言葉を用いて、言語形式だけでなく意味にも注意を払っている様子が伺えた。Mika は、MT の英文は必ず自分で手を加えてポストエディットする理由をこう説明した(抜粋5)。

抜粋5

(MTの英文を)そのまま使うと、それは自分の意見なのかって言われたら、ちょっと微妙だなんて思います。ちょっと罪悪感もあって (Mika, 141)。

情意エンゲージメントについては、5人の参加者がMT使用への前向きな態度を示した(表11)。

■表11: 情意エンゲージメント

中エンゲージメント	高エンゲージメント
- Google 翻訳はカメラ機能もあってすごく使いやすい (Aoi, 78) - MT がなかったら単位も取れなかったと思う (Saeko, 127)	- MT はめちゃくちゃいいと思う (Koki, 171) - MT があったら長い文で日記が書ける (Naomi, 195) - これからも MT に手助けしてもらおう (Taka, 184)

一方、高校時代に電子辞書を使って熱心に英語学習に取り組んでいたSaekoは、MTを使うことへの複雑な心情を表した(抜粋6)。

抜粋6

(オンライン授業で)緊張すると、中学生くらいの文も出てこないし、単語も忘れてしまう。(MTで)早く調べなきゃみたいになって。それで英語を身につけようっていうより、単位を取ることに必死になってしまって(107)。… 余計に(MTを)使うようになってしまいました(Saeko, 109)。

社会エンゲージメントについてはAoi, Taka, Yokoが友達からMTアプリを教えてもらっていた(表12)。特にAoiは、2つのアプリを使って自身の英語が留学生との会話で適切かどうか確かめる方法を寮の日本人の友達から教わっていた。実際に友達と一緒にMTを使う参加者はなく、MTは個人での使用に限られていた。

■表12: 社会エンゲージメント

中エンゲージメント	高エンゲージメント
・寮の友達から Google が出力した英文を DeepL で確かめることを習った(Aoi, 120)	・友達に Google を勧められた(Taka, 91) ・友達に Patago を勧められた(Yoko, 48)

主体性エンゲージメントが見られたのは、Naomi だけだった(表13)。

■表13: 主体性エンゲージメント

高エンゲージメント
英語で日記を書くとき、難しい文や知らない単語を Google 翻訳で調べてノートに書き留める(Naomi, 41)

Naomi は半年前から英語で日記をつけるようになった経緯をこう語った(抜粋7)。

抜粋7

去年の夏休みに大学の短期留学に参加したとき、同じセミナーにいた子がすごく英語が上手だったので、その子がしていることを参考に(英語日記を)始めました(29)。… 自分で(英語で)書けるけど、機械翻訳がなかったら、中学校で習った文法ばかりとか、簡単な文になってしまうと思います(Naomi, 193)。

インタビューの最後に、参加者に質問はないか尋ねたところ、高校時代は熱心に電子辞書を使っていたというMikaは、MTの使用について、こう投げかけた(抜粋8)。

抜粋8

今はもう電子辞書よりも機械翻訳がどんどん主流になっていくので、多分、辞書を持たない子のほうが将来、多くなってくると思うんですけど、そういう時に、学校って機械翻訳とどう付き合っていくのかなって思いますね(Mika, 184)。

このように、インタビューからは参加者がそれぞれの置かれた状況の中で、必要に応じて工夫しながらMTを使っていることが明らかになった。抜粋3と8のように、学習者もMTを使う経験を重ねながら、MTに対する自分なりのビリーフを確立させていくのかもしれない。

5 考察

5.1 MTエンゲージメント尺度の作成(調査1)

探索的因子分析ではMTエンゲージメントを構成する行動、プリエディット認知、ポストエディット認知、情意、社会の5つの因子が抽出された(RQ1)。これらの行動、認知、情意、社会エンゲージメント因子は先行研究による理論と一致した。主体性は、Yuasa & Takeuchi (in press)の質的データで示唆されたため尺度に取り入れたが、因子としては抽出されなかった。主体性エンゲージメントとは、学習意欲が高まる環境を作るための学習者の積極的な言動を指す(Reeve & Shin, 2020)。大学の英語教師の多くは学生のMT使用を支持していないこと(Yamada et al., 2021)が、学生のMT使用への積極的・主体的な行動を妨げているのかもしれない。そのことが、主体性因子が抽出されなかった要因になった可能性がある。主体性エンゲージメントの概念はまだ発展途上である(Patall, 2024)。構成概念を検証するためにはさらなる研究が必要である(Sinatra et al., 2015)。

MTエンゲージメント尺度の妥当性の検証では、尺度の信頼性と妥当性が満たされたが(RQ2)、MTエンゲージメントとMT自己効力感との相関はほぼ中程度で、併存的妥当性は弱い支持だった。これはMTの精度に対する学生のビリーフなど、他の要因が働いたからかもしれない。たとえばKoltovskaia(2020)は、人工知能(Artificial Intelligence; AI)を利用したライティングのフィードバック・ツール(Grammarly)への学習者のエンゲージメントを調べた。その結果、学習者がツールに依存するあまり、誤りを修正する際の方略使用が制限されたり、逆にツールへの懐疑心から方略の使用が増えたことを報告している。MTにおいても同様に、MTの出力精度への認識が、プリエディットやポストエディット方略の使用に影響した可能性がある。今回の調査で扱った要因以外の変数についても、今後、さらに調査すべきだろう。

5.2 MTエンゲージメントの量的・質的調査(調査2)

調査2では、CEFR A2レベルに相当するサンプルの大学生英語学習者をMTエンゲージメント尺度の変数(行動・プリエディット認知・ポストエディット認知・情意・社会)の類似グループに分類した。その結果、低エンゲージメント、中エンゲージメント、高エンゲージメントの3つのグループに分類された(RQ1)。グループ間の自己効力感の違いを比べたところ、高エンゲージメント・グループが高く、中・低エンゲージメント・グループは低かった(RQ2)。質問紙項目についての分散分析の結果(表4)とインタビューの結果から(RQ3)、学生は教師が想像する以上にMTの英文を確かめたり修正したりしていた。これはHellmich and Vinall (2023)の、フランス語やスペイン語を学ぶ米国の大学生を対象にした調査結果と一致している。特に全体の半数近く(45.6%)を占める高エンゲージメント・グループはMTを肯定的に捉え、より積極的に使っている可能性がある。このことはサウジアラビアや韓国、インドネシアなどのEFL学習者を対象とした先行研究でも報告されている(Almusharraf & Bailey, 2023; Briggs, 2018; Murtisari et al., 2019)。インタビューからは参加者が方略を使ってMTの英文を修正していることが伺えた。方略使用の程度や意識には差があるようだが、この結果は方略使用と自己効力感(Graham et al., 2020; Teng & Zhang, 2020)、エンゲージメントと自己効力感(Cai & Xing, 2023; Tsao, 2021)に正の相関があることと一致する。

中エンゲージメント・グループは認知エンゲージメントが比較的高く、参加者は方略使用について明確に語っていた。一方、Saekoがオンライン授業で孤立する中、MTを使っていたこと(抜粋2)や、MTに頼ることで英語が身につかないかもしれないという恐れ(抜粋6)、MikaのMT使用に対する「罪悪感」という言葉(抜粋5)は、このグループの情意と社会エンゲージメントの低さを説明している。その要因には、MTを使うことへの倫理的呵責(“ethical qualms”; Klekovkina & Denié-Higney, 2022, p.107)が影響し

ているのかもしれない。

5.3 教育的示唆

インタビューの参加者は全員、先行研究と同様に、MTの出力が不完全であることを認識しながらも (Vinall et al., 2023), 自分で書く英文よりも良い (Murtisari et al., 2024) からと課題にMTを使っていた。“MT-enabled L2 writing” (MTを活用したL2ライティング; Jiang et al., 2024, p. 1) が標準になる時代には、習熟度が低い学習者にとっては、成果物の正確さはもはやL2学習の成果とはみなされないだろう (Klekovkina & Denié-Higney, 2022)。MTの助けを借りながら、最終的なプロダクトにどのようにたどり着いたか、プロセスがより重要である。そこに教師による介入の余地がある。Lee (2021) はCEFR A2レベルの韓国人大学生に1学期間、MTを使ったL2ライティングを指導した。その結果、教師やクラスメイトの足場かけにより、メタ言語知識/言語への気づき (metalinguistic awareness) が向上し、L2ライティングへの自信が増した。

今回の質問紙調査では、社会エンゲージメントの低さが際立った。学習者同士で一緒にMTを使うというインタビュー参加者もいなかった。このことは、教師がMTの使用を認めていないため、MTをオープンに使うことへの倫理的呵責が影響しているのかもしれない。インタビューでは、Koki (202) の「合っているかわからないけど、時々時制を変える」や、Taka (121) の「一応、自分で英文を確認するが、それが正しいかどうかはわからない」などの表現 (表10) に、自身でMTの正確さを判断する自信のなさが表れていた。CEFR A2レベルの学習者は、教師の足場かけの元、協働でMTの英文を修正しながら適切な使い方を学ぶことが必要ではないか。この体験がMTを使って正しい英文が書けるという自己効力感を高め、さらにエンゲージメントを高める (Schunk & DiBenedetto, 2022)。学習者によるMTを使った言語使用についてのやりとりは、Svalberg (2009) が提唱する「言語へのエンゲージメント」 (“engagement with language”) を促進する可能性がある。Svalberg (2018) は「言語へのエンゲージメント」を言語への気づきが高まるプロセスと呼んでいる。言語への気づきはL2学習に不可欠である (Schmidt, 1990)。

Vinall et al. (2023) は、教師と学習者が対話しながらMT使用の方針を学習目的に合わせて柔軟に決めることを提案している。MTの不完全さは、教師と学習者がオープンに議論することで、学びの機会を創出することができる。習熟度が低い学習者にとっては、不完全なL2でやりとりをするよりも、MTを使ってより通じる英文へと磨き上げる活動は、有意義なL2学習になると確信している。自分がほんとうに伝えたい内容について、L1とL2を比較し、語彙や言語形式について議論することは、学習者の知的好奇心を刺激し、エンゲージメントを高めるだろう。

6 今後の課題

今回の調査には限界点もある。MTエンゲージメント尺度の作成では、男女比や学部の違いに配慮したが、一つの大学のみからの参加者を対象にしたので、バイアスがかかる可能性は否めない。他大学の学生も対象とすることで、妥当性を高めることができるだろう。因子として抽出されなかった主体性エンゲージメントに関しては、定義と項目の見直しも必要ではないかと考える。

インタビュー調査では、全体の3割弱を占める低エンゲージメント・グループに属する協力者がいなかったため、このグループの詳細を明らかにすることはできなかった。また7名の参加者からインタビューによりMTの使い方を聞き出したが、実際に使用している場面を確認したわけではない。より詳細に学生個人のMTの使い方を理解するためには、刺激着想法などの研究手法を併用することが望ましい (Hellmich & Vinall, 2023; Yuasa & Takeuchi, 2024)。

7 まとめ

本調査は、CEFR A2レベルに相当する大学生英語学習者のMT使用の現状を、エンゲージメントの枠組みに基づいてマクロとミクロの視点から明らかにした。質問紙調査では、MTエンゲージメント尺度の5つの因子(行動、プリエディット認知、ポストエディット認知、情意、社会)の強さにより、参加者は低・中・高のエンゲージメント・グループに分類された。中エンゲージメントと高エンゲージメント・グループを合わせたおよそ7割の参加者は、プリエディットやポストエディットなどの方略を用いてMTを使っていた。エンゲージメントが高いグループは自己効力感も高く、エンゲージメントが低いグループは自己効力感も低かった。すべてのグループで社会エンゲージメントが際立って低いことから、MTは個人での使用に限られているようだ。インタビューから、参加者がそれぞれの目的に合わせて、工夫しながらMTを使っていることがわかった。一方、MTの英文の正確さを十分に判断できていない可能性や、MTを使うことへの倫理的呵責も見えた。習熟度が低い学習者は、教師や学習者同士のやりとりを通して、MTを使いながら正しい英文へと磨いていく体験が必要である。このプロセスはL2学習に必要な言語への気づきを高めることができる。そこで初めて、MTは学習ツールとして利用することができるだろう。本調査がAI時代にふさわしい教授法を模索するきっかけになればと願う。

謝辞

本調査を実施するにあたり、貴重な機会とご支援をくださいました公益財団法人 日本英語検定協会と関係者の皆さま、選考委員の皆さまに心よりお礼申し上げます。特に研究助言者の寺内一先生には貴重なご助言をいただき、感謝申し上げます。また、今回の質問紙調査は、関西大学の多くの学生と先生方のご協力によって実現できました。さらにインタビューでは、7名の学生に機械翻訳についての貴重なご意見をいただきました。心より感謝申し上げます。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

- Almusharraf, A., & Bailey, D. (2023). Machine translation in language acquisition: A study on EFL students' perceptions and practices in Saudi Arabia and South Korea. *Journal of Computer Assisted Learning*, 39(6), 1988-2003. <https://doi.org/10.1111/jcal.12857>
- Bai, B., & Wang, J. (2023). The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research*, 27(1), 207-228. <https://doi.org/https://doi.org/10.1177/1362168820933190>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman and Company
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Front Public Health*, 6. <https://doi.org/10.3389/fpubh.2018.00149>
- Bowker, L., & Buitrago, J. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing.
- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal*, 14(1), 3-24. <https://doi.org/10.29140/jaltcall.v14n1.221>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
- Cai, Y., & Xing, K. (2023). Examining the mediation of engagement between self-efficacy and language achievement. *Journal of Multilingual and Multicultural Development*, 1-13. <https://doi.org/10.1080/01434632.2023.2217801>
- Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96, 102408. <https://doi.org/10.1016/j.system.2020.102408>
- Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal*, 100(S1)(Suppl. 1), 64-80. <https://doi.org/10.1111/modl.12302>
- Chung, E. S. (2020). The effect of L2 proficiency on post-editing machine translated texts. *The Journal of AsiaTEFL*, 17(1), 182-

引用文献

193. <https://doi.org/10.18823/asiatefl.2020.17.1.11.182>
- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). Sage Publications.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Ducar, C., & Schocket, D. H. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google Translate. *Foreign Language Annals*, 51(4), 779-795. <https://doi.org/10.1111/flan.12366>
- Gally, T. (2019). The implications of machine translation for English education in Japan. *JACETSIG-ELE Journal* 6(2), 1-14. <https://ci.nii.ac.jp/naid/40021993334/>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning and Technology*, 26(2), 5-24. <https://doi.org/10.125/73474>
- Goetze, J., & Driver, M. (2022). Is learning really just believing? A meta-analysis of self-efficacy and achievement in SLA. *Studies in Second Language Learning and Teaching*, 12(2), 233-259. <https://doi.org/10.14746/ssllt.2022.12.2.4>
- Graham, S. (2022). Self-efficacy and language learning: What it is and what it isn't. *The Language Learning Journal*, 50(2), 186-207. <https://doi.org/10.1080/09571736.2022.2045679>
- Graham, S., Woore, R., Porter, A., Courtney, L., & Savory, C. (2020). Navigating the challenges of L2 reading: Self-efficacy, self-regulatory reading strategies, and learner profiles. *The Modern Language Journal*, 104(4), 693-714. <https://doi.org/10.1111/modl.12670>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hellmich, E. A., & Vinall, K. (2023). Student use and instructor beliefs: Machine translation in language education. *Language Learning and Technology*, 27(1), 1-27. <https://hdl.handle.net/10125/73525>
- Hiver, P., Al-Hoorie, A. H., & Mercer, S. (2021). *Student engagement in the language classroom*. Multilingual Matters.
- 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>
- In'nami, Y., & Koizumi, R. (2011). Structural equation modeling in language testing and learning research: A review. *Language Assessment Quarterly*, 8(3), 250-276. <https://doi.org/10.1080/15434303.2011.582203>
- Jiang, L., Yu, R., & Zhao, Y. (2024). Theoretical perspectives and factors influencing machine translation use in L2 writing: A scoping review. *Journal of Second Language Writing*, 64, 101099. <https://doi.org/10.1016/j.jslw.2024.101099>
- Kern, R. (2024). Twenty-first century technologies and language education: Charting a path forward. *The Modern Language Journal*, 108(2), 515-533. <https://doi.org/10.1111/modl.12924>
- Klekovkina, V., & Denié-Higney, L. (2022). Machine translation: Friend or foe in the language classroom? *L2 Journal*, 14(1), 105-135. <https://doi.org/10.5070/I214151723>
- Klimova, B., Pikhart, M., Benites, A. D., Lehr, C., & Sanchez-Stockhammer, C. (2022). Neural machine translation in foreign language teaching and learning: A systematic review. *Education and Information Technologies*, 28(1), 663-682. <https://doi.org/10.1007/s10639-022-11194-2>
- Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google translate in academic writing courses? *The EUROCALL Review*, 26(2), 50-57. <https://doi.org/10.4995/eurocall.2018.10140>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450. <https://doi.org/10.1016/j.asw.2020.100450>
- Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157-175. <https://doi.org/10.1080/09588221.2018.1553186>
- Lee, Y.-J. (2021). Still taboo? Using machine translation for low-level EFL writers. *ELT Journal*, 75(4), 432-441. <https://doi.org/10.1093/elt/ccab018>
- Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: an international journal*, 46(3), 517-528. <https://doi.org/10.2224/sbp.7054>
- Liu, M., Noordin, N., Ismail, L., & Abdrahim, N. A. (2023). Relationship between student engagement and academic achievement in college English education for non-English majors in China. *International Journal of Learning, Teaching and Educational Research*, 22(8), 203-232. <https://doi.org/10.26803/ijlter.22.8.12>
- Lo, S. (2023). Neural machine translation in EFL classrooms: Learners' vocabulary improvement, immediate vocabulary retention and delayed vocabulary retention. *Computer Assisted Language Learning*, 1-20. <https://doi.org/10.1080/09588221.2023.2207603>
- Mackey, A., & Gass, S. M. (2022). *Second language research: Methodology and design* (3rd ed.). Routledge.
- Mercer, S. (2019). Language learner engagement: Setting the scene. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 643-660). Springer.
- Mercer, S., & Dörnyei, Z. (2020). *Engaging language learners in contemporary classrooms*. Cambridge University Press.
- Murtisari, E. T., Kristianto, A. K., & Bonar, G. (2024). Self-directed use of machine translation among language learners: Does it lead to disruptive L2 avoidance? *Foreign Language Annals*, 1-24. <https://doi.org/10.1111/flan.12768>
- Murtisari, E. T., Widiningrum, R., Branata, J., & Susanto, R. D. (2019). Google Translate in language learning: Indonesian EFL

引用文献

- students' attitudes. *The Journal of Asia TEFL*, 16(3), 978-986. <https://doi.org/10.18823/asiatefl.2019.16.3.14.978>
- Negishi, M., Takada, T., & Tono, Y. (2013). A progress report on the development of the CEFR-J. In E. D. Galaczi & C. J. Weir (Eds.), *Exploring language frameworks: Proceedings of the ALTE Kraków conference, July 2011* (pp. 135-163). Cambridge University Press.
- O'Neill, E. M. (2019). Online translator, dictionary, and search engine use among L2 students. *CALL-EJ*, 20(1), 154-177. <https://old.callej.org/journal/20-1/O'Neill2019.pdf>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578. <https://doi.org/10.3102/0034654306004543>
- Patall, E. A. (2024). Agentic engagement: Transcending passive motivation. *Motivation Science*, 10(3), 222-233. <https://doi.org/10.1037/mot0000332>
- Reeve, J., & Shin, S. H. (2020). How teachers can support students' agentic engagement. *Theory Into Practice*, 59(2), 150-161. <https://doi.org/10.1080/00405841.2019.1702451>
- Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, 36(4), 257-267. <https://doi.org/10.1016/j.cedpsych.2011.05.002>
- Reschly, A. L., & Christenson, S. (2022). *Handbook of research on student engagement* (2nd ed.). Springer.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158. <https://doi.org/10.1093/applin/11.2.129>
- Schunk, D. H., & DiBenedetto, M. K. (2022). Self-efficacy and engaged learners. In A. L. Reschly & S. Christenson (Eds.), *Handbook of research on student engagement* (2nd ed., pp. 155-170). Springer.
- Shin, D., & Chon, Y. V. (2023). Second language learners' post-editing strategies for machine translation errors. *Language Learning and Technology*, 27(1), 1-25. <https://hdl.handle.net/10125/73523>
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, 50(1), 1-13. <https://doi.org/10.1080/00461520.2014.1002924>
- Staples, S., & Biber, D. (2015). Cluster analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 243-274). Routledge.
- Stapleton, P., & Kin, B. L. K. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes*, 56, 18-34. <https://doi.org/10.1016/j.esp.2019.07.001>
- Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, 90, 102221. <https://doi.org/10.1016/j.system.2020.102221>
- Svalberg, A. M. L. (2009). Engagement with language: Interrogating a construct. *Language Awareness*, 18(3-4), 242-258. <https://doi.org/10.1080/09658410903197264>
- Svalberg, A. M. L. (2018). Researching language engagement: Current trends and future directions. *Language Awareness*, 27(1-2), 21-39. <https://doi.org/10.1080/09658416.2017.1406490>
- Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *Journal of Second Language Writing*, 48, 100701. <https://doi.org/10.1016/j.jslw.2019.100701>
- Tsai, S.-C. (2019). Using google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 32(5-6), 510-526. <https://doi.org/10.1080/09588221.2018.1527361>
- Tsao, J.-J. (2021). Effects of EFL learners' L2 writing self-efficacy on engagement with written corrective feedback. *The Asia-Pacific Education Researcher*, 30(6), 575-584. <https://doi.org/10.1007/s40299-021-00591-9>
- Vinall, K., Wen, W., & Hellmich, E. A. (2023). Investigating L2 writers, uses of machine translation and other online tools. *Foreign Language Annals*, 57(2), 499-526. <https://doi.org/10.1111/flan.12733>
- Wang, C., & Sun, T. (2020). Relationship between self-efficacy and language proficiency: A meta-analysis. *System*, 95, 102366. <https://doi.org/10.1016/j.system.2020.102366>
- Williams, M., & Burden, R. L. (1997). *Psychology for language teachers: A social constructivist approach*. Cambridge University Press.
- Yamada, M., Langliz, H., Oda, T., Morita, T., Tamura, H., Hiraoka, Y., & Irie, T. (2021). A preliminary survey on the use of machine translation in English education at Japanese universities. *Invitation to Interpreting and Translation Studies*, 23, 139-156. https://doi.org/http://honyakukenyu.sakura.ne.jp/shotai_vol23/No_23_007-Yamada_et-al.pdf
- Yuasa, M., & Takeuchi, O. (2022). Strategic use of Google Translate by Japanese university students of English [Conference presentation]. *Situating Strategy Use: Strategic Learning in an uncertain world (SSU4)*, Wellington, New Zealand.
- Yuasa, M., & Takeuchi, O. (2024). Stimulated recall interviews using digital technology in L2 research. In K. Sadeghi (Ed.), *Handbook of technological advances in researching language learning*. Routledge.
- Yuasa, M., & Takeuchi, O. (in press). Strategic use of machine translation: A case study of Japanese EFL university students.
- Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). Cambridge University Press.

資料1：半構造化インタビューの質問項目

Part 1: 英語学習について

1. 英語学習は好きですか？
2. その理由は？
3. あなたの習熟度はどれくらいだと思いますか？
4. 英検やTOEICなどの英語の資格試験を受験したことはありますか？

Part 2: ライティング活動について

1. 英語のライティングは好きですか？
2. 高校の英語の授業でライティングの指導は受けましたか？
3. どんな内容でどのくらいの指導がありましたか？
4. ライティングのどこが難しいと思いますか？

Part 3: MTを使う背景や目的, MTの精度について

1. どのMT(アプリ)を使っていますか？
2. いつ頃からMTを使い始めましたか？
3. MTを使うようになったきっかけは何だと思いますか？
4. MTを使う理由や目的は何ですか？
5. MTのどんなところが良いと思いますか？
6. MTのどんなところが良くないと思いますか？
8. MTの訳は正確だと思いますか？

Part 4: MTの使い方について

1. MTを使うとき, 初めから日本語を入力しますか?自分で英文を書いてから使いますか？
2. 入力する日本語を直しますか？
3. MTが出力した英文を自分で見直していますか？
4. MTが出力した英文を自分で修正しますか？
5. 修正するときに, MT以外に辞書などを参照していますか？
6. どんなもの(辞書・ウェブサイトなど)を参照しますか？
7. MTの使い方を誰かに教えてもらったことはありますか？
8. 今後もMTを使い続けたいですか？
9. ChatGPTなどの, 他のAIツールを使ったことがありますか？

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

吉澤孝幸先生の実践研究「用法基盤モデルに基づいたスピーキング指導が中学生の『即興力』育成に及ぼす効果の検証」は、日常の授業を参観しているような研究報告であり、全国の小中高の英語教師にとって、日常の英語授業における研究テーマを設定して授業を行う際の視点として、大変参考になる実践報告である。研究テーマの設定・先行研究の活用・研究目的・研究方法・実践研究における評価基準の設定方法・評価の分析方法・評価結果および考察は大変参考になります。また、引用文献の幅広さに敬服いたします。是非、全国の小中高の英語教師の皆さんには、日常の授業に研究目標を設定して授業を行ってほしい。

研究者の吉澤孝幸先生には、今回の「用法基盤モデルに基づいたスピーキング指導が中学生の『即興力』育成に及ぼす効果の検証」のテーマを盛り込んだ公開授業を積極的に行い、多くの英語教師の模範になることを期待いたします。



名誉顧問 小池 生夫

本年で3回目の研究発表についての所感を述べる。定年齢に達した者は、このポストを比較的若い研究指導者に譲り、より一層の知的指導を期待するのが一般であろう。今回も彼らは、その責をその講評において十分果たしていると思い、以下に、若干の特徴を述べる。

本年の研究者は、研究部門では大学院生が占め、実践部門では高校、中学校の教員、調査部門では中学校の教員と大学院生であった。しかしその裏では、実践と調査各部門の教員の方々は、実は大学院で研究を積んでおり、そこで培った知識、能力を背景に研究を発表していて、その培った実力を今も保っている。これは最近数年の傾向である。実践部門では、生徒を相手に発話の研究ができる環境を十二分に活用して、debate、discussionの対応、養成、発表の技能育成をしている。調査部門では、機械翻訳を学習ツールとして役立たせるなど、発話能力の育成に向けた新しい開発法を見せている。

研究者の出身大学院については、伝統的に「英検」研究助成に力を入れてきた筑波大学大学院が引き続き存在感を示す一方、新しい大学院関係者も増加しており、その裾野が広がりとつあることを感ずる。また、内容面では、指導要領の改訂に伴う新分野への挑戦が行われており、いかにも若さが目立つ。特に機械翻訳など、新知識を導入して発話能力の向上に役立たせようとする新しい傾向が注目される。今回もchallengingな傾向が現れているのは、いかにも若い研究者らしい新鮮味があって、喜ばしい。

名誉顧問 村木 英治

この第36回「英検」研究助成報告のゲラ刷りを受けとった週に、EF Education Firstが発表した英語圏外の国々の「英語能力指数」の記事を読んだ。2024年版の指数によると、日本は韓国や中国にも負け、その低下に歯止めがかかっていないという。複雑多様な英語力をsingle indexで表すことは無理があろうが、それは現実の何かを指し示しているには違いない。自国言語のみで英語ができない社会人でも、不利な境遇に陥ることのない我が国は幸福であるのかもしれない。しかし、ほんの少し立ち上がり飛び上がりたい人間にとって、英語力は不可欠な能力だ。英語力はより国際化を迎える我が国にとって決して贅沢な能力ではない。そんなことを考えていると、本誌の小出氏の「日本語母語英語学習者が使用する定型表現の分析」の論文に関心がとまった。本号では、speakingとかwritingなどの能動的な英語行動にフォーカスした研究論文が多い気がする。能動的な英語行動であればあるほど、日本人特有の英語が滲みでてくる。英検協会が出版する日本語母語話者の英語についてのこれらの研究論文は、ますます貴重なものになるであろう。また本誌に掲載されている論文を読み進めていくうちに、「自立した、あるいは自律した英語学習者」という共通概念が浮かびあがってきた。それは高い「英語能力指数」と同じものなのだろうか。本誌の研究者たちと考えていきたい。

名誉顧問 和田 稔

第37回の「英検」研究助成への申請者の総数は大幅に増加したという。加えて、本誌第36回でも、「研究部門」「実践部門」「調査部門」の入選者のバランスがよいのは喜ばしいことである。

申請者が増加した理由を入選者の研究テーマだけから判断するのは慎重であるべきと思うが、やはり学習指導要領の改訂(小学校 平成29年, 中学校 平成29年, 高等学校 平成30年告示)が影響していると思われる。入選者以外の研究テーマを分析することも大切であろう。

さらには、「実践部門」の入選が4件であり、高等学校対象が3件, 中学校対象が1件であることも妥当と思われる。改訂された学習指導要領では高等学校の履修科目と科目の履修構成の点において、多様化している。つまり、義務教育段階では、全国的に一定の教育水準を確保することが目標であるが、高等学校においては多様な学習を確保することがねらいとなる。このような理由から高等学校における科目が多様化し、深化するのは当然であろう。しかし、このような英語学習の位置づけの違いはあるが、義務教育段階の小・中学校での英語指導にも研究すべき課題があるのは当然であろう。

それぞれの論文は研究テーマの選択、研究手法などの点で、参考になることは多い。これから論文をもとに、研究成果の共有が進むことを期待したい。

ISSN 1348-7949

EIKEN BULLETIN vol. 36 2024

非売品

2025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>

印刷・製本／日新印刷株式会社

無断複製・転載を禁じます。 ©2025 公益財団法人 日本英語検定協会