

EIKEN BULLETIN

vol. 32 2020

第32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

公益財団法人 日本英語検定協会

第32回「英検」研究助成 選考委員

(役職は委嘱当時、*印は専門選考委員) ※五十音順

- * 池田 央 立教大学 名誉教授
- * 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
- * 大友 賢二 筑波大学 名誉教授
- * 長 勝彦 一般財団法人 語学教育研究所 参与
- 川越 豊彦 全日本中学校長会 会長
- * 小池 生夫 慶應義塾大学 名誉教授・明海大学 名誉教授
- * 小泉 利恵 順天堂大学 准教授
- 鈴木 真人 全国英語教育研究団体連合会 会長
- * 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
- * 村木 英治 東北大学 名誉教授
- 山田 杉子 全国高等学校長協会 国際教育交流委員会 委員長
- * 吉田 研作 上智大学 特別招聘教授・言語教育研究センター長
- * 和田 稔 明海大学 名誉教授

はじめに

上智大学 特別招聘教授・言語教育研究センター長
吉田 研作

はじめに、毎年数多くの研究の応募があるが、大学院生、現場の先生方の問題意識の高さと研究への熱意に敬意を表したい。

さて、例年、「英検」研究助成の申請書を見ているが、その経験から少し述べることにする。

最初に、大学院生の研究計画について気づいたことを述べることにする。大学院生の研究論文は、研究テーマの設定、研究方法、分析方法という点で論理的にもしっかりと計画されているものが多い。中には、確かに素晴らしい内容のものもある。しかし、必ずしもすべてがそうだというわけではない。中には、研究計画はしっかりとしているが、最初から結果が十分予想されるようなテーマのものがあったり、逆に、テーマは面白いが、被験者の選定や研究のための実験材料が、研究テーマとして設定された課題に適していないのではないか、と思われるものがある。どんなに素晴らしい研究計画でも、それを検証するための具体的な方法論(被験者、実験材料等)が伴わなければ良い成果は得られないだろう。本研究助成は、実践的な英語教育とその評価に寄与することが目的になっているが、そのための研究方法が採用されているかどうか今一度検討していただきたい。

他方、英語教育の現場で日ごろから問題意識を持ち、本研究助成に応募された先生方の中には、問題意識は良いが、それを研究や調査のテーマとしてどのように検証あるいは調査すれば良いかが良くわからないまま応募される方がおられるようと思う。調査や研究には、それなりの方法があるので、それについてもう少し調べていただきたいと思うことが良くある。また、ものによっては先生が置かれている個別の状況にあまりにも限定されすぎていて、提案されている研究の汎用性について疑問を持たざるを得ないものもある。

「英検」研究助成の特徴は、研究、調査、実践等、色々な分野の研究を認めているところにある。しかし、理論的研究は現場の英語教育にどのように適応できるのか、また、実践的研究は、でてきた結果が理論的にも納得できる汎用性のあるものである必要がある。

これから英語教育は益々多様化していくだろう。そのためにも様々な研究に基づいたより良い英語教育の姿を模索していく必要があることを忘れないでいただきたい。

Contents

● はじめに

上智大学 特別招聘教授・言語教育研究センター長 吉田 研作 3

● 報告別講評 専門選考委員 7

A 研究部門

ライティングタスク(技能独立型vs.統合型)が発表語彙とその測定に 与える影響: TAALESによる語彙の洗練性分析を基に

研究者: 茨城県／筑波大学大学院 在籍 小室 竜也 13

自動英文解析ツールを用いた英作文採点の妥当性検証: Coh-MetrixとText Inspectorの指標に基づいて

研究者: 茨城県／筑波大学大学院 在籍 佐々木 大和 31

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 —

研究者: 岐阜県／岐阜県立 関高等学校 教諭 福田 陽子 44
申請時: 名古屋大学大学院 在籍

多様な言葉の働きを引き出すタスクタイプの組み合わせ — 言語機能分析による提案 —

研究者: 茨城県／筑波大学大学院 在籍 前田 啓貴 68

B 実践部門

ループリックの事前提示がライティングパフォーマンスに与える影響

研究者：東京都／中央大学附属中学校・高等学校 教諭 岩本 祐樹 …… 85

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

研究者：東京都／青山学院 高等部 特別教諭 江下 陣 …… 108
申請時：福岡県立 小倉南高等学校 教諭

AIを活用したライティング能力の育成

— CEFR B1からB2へのレベルアップを狙って —

研究者：新潟県／新潟県立 津南中等教育学校 教諭 松井 市子 …… 125

Pre-readingでのトップダウンアプローチとPost-readingの統語処理活動を取り入れた高校生の多読授業実践

研究者：東京都／東京都立 国分寺高等学校 教諭 三上 洋介 …… 144
申請時：東京都立 大田桜台高等学校 教諭

高等学校での内容言語統合型学習(CLIL)による他教科と連携した授業の実践と効果の検証

研究者：栃木県／栃木県立 宇都宮南高等学校 教諭 渡邊 聰代 …… 173

C 調査部門

学習者の視点から見た英語による専門科目(EMI)に必要な英語力のニーズ分析

研究代表者：東京都／早稲田大学大学院 在籍 太原 達朗 …… 190

日本人大学生の英語学習における自己調整学習能力尺度の開発：英語資格試験に向けた自主学習に焦点を当てて

研究者：東京都／立教大学大学院 在籍 福田 晶子 …… 218

- 第32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	2019年2月1日～4月5日
選考	4月12日～6月22日
助成金贈呈式	7月13日
研究期間	7月～2020年5月上旬
報告書提出	2020年5月10日

報告別講評

A. 研究部門・報告 I

寺内 一

**ライティングタスク
(技能独立型vs.統合型) が
発表語彙とその測定に与える影響：
TAALESによる語彙の洗練性分析を基に**

【研究者：小室 竜也】

本研究は、日本人英語学習者を対象に2種類のライティングタスクを実施し、タスクの違いが作文内で用いられている発表語彙にどのような影響を与えていているのかを検証することを目的として2つの調査を行ったライティングに関する研究である。ご存知のように、語彙知識は、それが実際の言語使用の際にどのように活用されるかという観点から、受容語彙と発表語彙とに分類される。そして、受容語彙と発表語彙はそれぞれが独立したものではなく、受容語彙から発表語彙に連続体をなしていくように語彙知識として習得されていくとのが一般的に言われている。本研究は、技能独立型(特定のトピックについて自分の意見を述べる形式)と技能独立型および技能統合型(トピックに関する英文を読解後、自分の意見を書く形式)の2種類の異なるタスクを通してそれぞれの被検者が使用している発表語彙の特徴にどのようなものがあるのかを把握しようとしたものである。

学習指導要領の5領域において、ライティングの指導と評価は重要な課題であり、本研究がライティングのタスクと発表語彙の関連性を検証した結果を丁寧にまとめていることは高く評価したい。特に、結論部分でも触れているように、現在の中高(中等教育)での連語の指導が受容において明示的に、発表において暗示的に指導されることが多いため、連語表現が発表語彙として未発達であったとの指摘は、中高だけでなく、大学でのアカデミックライティング指導に応用できるであろうという提言は、大いに参考にしていくべきであろう。

A. 研究部門・報告 II

村木 英治

**自動英文解析ツールを用いた
英作文採点の妥当性検証：
Coh-Metrix と Text Inspector の
指標に基づいて**

【研究者：佐々木 大和】

現に与えられているデータとそこから generate できる仮説の限界性が読みながら自然と湧き上がってくる。それはこの論文の欠点ではけっしてなく、むしろこの次の研究段階の方向性を exciting に指示してくれているこの論文の優れた点のような気がする。それは著者も幾度も指摘しているように英作文のトピックとそれに伴う特徴、あるいは subscale 得点の optimal な weights との関係である。わたしも過去に英語の native が書く英作文とその type について研究したことがある。因子分析法にかけると英作文における persuasive, narrative, informative といった文章のタイプによりその特徴にかなりきわだった違いがある。日本人が書く英作文について、type や topic の面からさらなる研究が進むことは望ましい。優れた論文とは重要な問題点を明瞭に指示し、かつその点についてのさらなる展開につなげができる道を用意してくれているものであるならば、佐々木氏の本論文は明らかに成功している。

A. 研究部門・報告Ⅲ

寺内 一

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—

【研究者：福田 陽子】

本研究は、ライティングにおける文章（エッセイ）の一貫性と結束性に関して量的・質的観点から分析を行うことにより、学習者がエッセイを書く際の言語的特徴と、それらが完成したエッセイに対してどのような影響を及ぼしているかを見出した非常に興味深い研究である。

本研究の特筆すべき点は以下の2点に集約できる。1点目は、日本人英語学習者と英語母語話者が作成したエッセイを比較検討した結果、日本人学習者は主語の人称代名詞と、文頭・節間ににおけるつなぎ語の過剰使用があることにより、論理の飛躍や因果関係の欠如が起きてしまうことが判明したのである。もう1点は、日本人学習者のエッセイ内の一貫性・結束性とエッセイのスコアとの有意な相関があり、特につなぎ語・指示代名詞がエッセイの不自然さを左右する大きな要因であることが明らかになったのである。

福田氏自身は「結論と今後の課題」において、教育的示唆として、一貫性・結束性の要(かなめ)となる「つなぎ」語や指示代名詞の重要性を学習者に伝える方法等を提示しているが、日本人学習者が日本語を含めて英語で「エッセイ」を書くということに対して、それまでの学習過程においてどのような指導を受けていたのかに大きく影響を受けている可能性があるのかもしれない。本研究そのものとは直接は関係しないものの、今後の研究の方向性のひとつとして参考にしていただければ幸いである。

A. 研究部門・報告Ⅳ

和田 稔

多様な言葉の働きを引き出す
タスクタイプの組み合わせ

—言語機能分析による提案—

【研究者：前田 啓貴】

本研究は近年の日本の学校英語教育で重要視されている「タスク」がどのような「言語機能」（言語の働き）を引き出すか、を研究した成果をまとめたものであり、時機を得た研究である。具体的には、6種類のタスク（ナレーション、ジグゾー、情報ギャップ、問題解決、意思決定、意見交換）別に、どのような言語の働きが、どのような頻度で使われているか詳細に調べている。たとえば、問題解決タイプが、最も多くの言語機能を引き出していること、言語機能を引き出すことが最も少なかったタスクはナレーションであること、などが実証されている。このような研究結果は日本の学校英語教育に求められているタスクを学習段階別に設定するに当たって貴重な情報を提供している。その一例として、英語の熟達度の低い学習者にはナレーションタイプのタスクや意見交換タスクが有効であり、熟達度の高い学習者には、情報ギャップ型や問題解決型タスクが有効であることが指摘されている。

しかし本研究は言語機能（言語の働き）の使用という点に焦点化したものであり、学校英語教育においてコミュニケーション能力（特に、話すことのコミュニケーション能力）の向上のためのタスクを効果的に活用するには様々な観点からの研究が今後必要であろう。このことを意識した記述が本論文に散見する。たとえば、「同じタスクタイプの中でも異なるタスクを用いて調査する必要がある」「協力者は熟達度が近い他の協力者と2人1組のペア」など。今後、これらの点を踏まえて本研究を深めることを期待したい。

B. 実践部門・報告 I

小泉 利恵

**ルーブリックの事前提示が
ライティングパフォーマンスに
与える影響**

【研究者：岩本 祐樹】

本研究は、ライティング評価のためのルーブリックをどのように指導で活用するかを考える上で重要な研究である。ライティングや生徒の捉え方を分析し、事前提示を行った実験群と行わなかつた統制群の間で大きな違いは見られなかつたが、その結果を先行研究と比較しながら丁寧に解釈している。

この研究で特に工夫されている点は、実験デザインである。実験群とともに統制群を設けたことで、ルーブリックの事前提示の影響をより詳しく調べることができている。学校内で統制群を設けることに抵抗感があるかもしれないが、実験後に、統制群の生徒にも統制群で行った良い効果があると考えられる指導を行えば、統制群の生徒が不利益を被ることはあまりない。また、複数の課題を使うとその難易度の違いの影響や、複数回テストを受けることによる順序効果も考えられるが、その影響を相殺するために、カウンターバランスを行っている。加えて、採点者を研究者とは別の教員に依頼し、採点する答案が実験群と統制群のどちらに属するかが分からぬよう形で採点している。ライティングやスピーキングの研究では他の教員に採点依頼がなかなか難しく、研究者自身が行うことも多く、自分の意図する仮説に近くなるような採点をしないように注意が必要だが、本来は本研究で行ったような形が望ましい。今後、本研究の後半で述べたような、ルーブリック提示・使用方法の改善や、自己評価・相互評価と組み合わせた指導を行い、さらに実践や研究を深めていってほしい。

B. 実践部門・報告 II

和泉 伸一

**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**

【研究者：江下 隣】

スピーキングの指導で生徒へいかにフィードバックするかということで悩まれている英語教師は少なくない。個別にフィードバックするには時間がかかるし、行ったとしても実際に時間対効果がどこまであるのか疑問に思うことも少なくない。せっかくの「親心」からのフィードバックが生徒のやる気を削いでしまったりすることも珍しくない。そういう際に有効な手だての一つとして、モデル文の活用がある。

従来の英語授業では、モデル文を最初に与えて、それに沿った形で自らのスピーチを考えて行うといった、「モデル提示」→「生徒の発話」といったパターンが多かった。一方、この論文で江下氏が提案しているのは、「生徒の発話」→「モデル提示」→「発話の再トライ」という流れである。いわば、「親心」から先回りして生徒に英語を教えようとするのではなく、「可愛い子には旅をさせよ」の気持ちで、まず生徒に挑戦させた後で、モデル文に触れさせ学びの機会を与え、再トライさせてみるといったものである。この手法の一つの大きな利点は、最初にアウトプットさせることで生徒自身が「学びの必要性」に気づくことである。また、そういった必要感が高まった上でモデルに触れさせることで、「栄養の吸収率」を高めることが期待される。同時に、モデル文から学ぶことは生徒自身が決めるので、学習者主体の学びを促すことができる。

これまでの研究でライティングにおいてこういった手法の効果は確認されてきたが、江下氏の研究で、それがスピーキングの指導でも当てはまることが示された。解決が必要な課題はまだまだあるが、今後の実践と研究がより期待される分野である。

B. 実践部門・報告Ⅲ

小泉 利恵

AIを活用した ライティング能力の育成

— CEFR B1 から B2 へのレベルアップを狙って —

【研究者：松井 市子】

B. 実践部門・報告Ⅳ

長 勝彦

Pre-reading での トップダウンアプローチと Post-reading の統語処理活動を 取り入れた高校生の多読授業実践

【研究者：三上 洋介】

本研究は、AI搭載型の Write and Improve を用いてライティングの指導と評価を高校1年生に行った結果、CEFR レベルの A2 から B1 への向上が見られたことを報告している。この研究では、4回分の定期テスト、5回分の授業中に行なった練習タスク、さらに4回分の模擬試験、2年間4回分の民間英語検定試験の結果を、様々な角度から検証している。本文で述べられているように、これらのテストは難易度が調整されていないものもあるため、点数が伸びたとしても力が伸びたのかタスクが易しかったのかが分からぬという欠点はある。しかし、長期的なパターンで見ていくことで全体的な傾向を見つけ、指導の効果や改善点を見つけていくことは可能であり、それを試みた本研究は非常に価値がある。

また、民間英語検定試験で伸びが見られた生徒とそうでなかった生徒を選び、その生徒のスコアがどのような変遷をたどり、Write and Improve や他の教材を生徒がどのように捉えて使っていたかを、上記のテストに加えてアンケート結果を用いて述べている。それにより、AI を使ったライティング指導の強みや弱点が見えており、今後どのように、強みをさらに活かし、弱点を全体の英語指導の中で中和させていくかのヒントが得られている。

松井氏が用いた手法は、指導の効果検証のための研究を日々の授業の中で進めるモデルとなるだけでなく、教員や学年、学校、また地域の PDCA サイクルを回して教育・授業改善を行う際の参考にもなるだろう。

研究者の学校では、実社会で使える英語力とビジネス知識の習得を目指している。生徒は1年生から、「使える英語を楽しく身に付けさせる」ことを目標に、全学的なカリキュラムとして必修授業に於ける多読・多聴学習を英語科全教師が取り組んでいる。

三上氏は、校務、部活動指導等で大変忙しい中、多読指導に関する多くについて、実に周到な準備及び、きめ細かい実践研究に取り組み、研究報告を完成した。三上氏の目の前の生徒が興味を持って多読に取り組む姿が手に取るようにわかる実践研究報告である。英語教師にとって多読指導の参考となる研究報告である。

「多読授業の効果」、「図書室で行う多読授業」、「CALL 教室で行う多読授業」、「多読多聴授業のツール(多読多聴手帳:多読多聴地図)の活用」、「多読多聴学習の導入用必読書リスト」、「推奨読書リスト」、「多読導入時のオリエンテーション」、「多読指導」、等々、実に具体的に報告書にまとめていいる。多読多聴学習の効果を測定し、統計的に処理して実証もしている。指導過程において、多読学習支援ワークシートを開発して、プレリーディングに於けるトップダウン処理を促して背景知識を活性化させ、生徒の読書意欲を引き出し、授業50分間集中して生徒は多読に取り組むようになった。巻末の参考資料(Worksheet、授業風景の写真等)、参考文献等も、英語教師にとって、今後多読指導を自分の授業に取り入れようとする際、大変参考になる。

「三上氏の授業を受けた生徒は、何と幸せであったことか」と、率直に思った。

B. 実践部門・報告V

和泉 伸一

**高等学校での
内容言語統合型学習(CLIL)による
他教科と連携した授業の実践と
効果の検証**

【研究者：渡邊 聰代】

C. 調査部門・報告I

和田 稔

**学習者の視点から見た
英語による専門科目(EMI)に必要な
英語力のニーズ分析**

【研究者：太原 達朗】

英語授業は語学の授業であるからといって、言葉の学習だけをしていればいいというものではない。言葉はそこに伝えるに値する内容があるからこそ重要であることを考えたとき、いかに言葉と内容学習を統合して教えていくかが重要になってくる。内実ある英語授業を行う上で、英語科の教員にとって問題となり得るのは、内容面での充実であろう。そこで昨今注目を集めているのが、他教科との連携である。専門性が異なり、また互いに忙しい教員同士で、どうしたらそういった連携が成しうるのか。どのような連携が可能で、その結果創出される授業の姿とはどういったものなのか。そして、それを生徒はどう受け止めるのだろうか。

こういった様々な疑問に実証研究の手法を用いて答えようとした試みが、渡邊氏の研究である。授業準備のあり方から始まり、授業観察、生徒への質問紙調査と連携教員へのインタビュー及び英語教員自身の振り返りといった様々なデータを収集して、プロセス指向の質的な研究を行っている。

筆者は渡邊氏が行った現代社会の教諭との連携授業を実際に見る機会があった。そこで印象的であった点は少なくとも三つある。(1)社会科の教諭が絵や写真を駆使して懸命に英語で生徒に語りかけている姿。(2)それを英語教員が適宜巧みに支援しつつ、生徒の英語学習に寄与しようとしている姿。そして(3)そこで展開される内容豊かな英語授業に生徒が鋭く反応して、同時に教員の挑戦する姿に触発されて、一生懸命に英語で討論しようとしている姿である。詳しい研究内容は渡邊氏の論文に書かれている通りであるが、教員が果敢に挑戦する中で深みと人間味のある授業が生まれ、そこに生徒と教員共々を交えた「学びの共同体」が自然と生まれてくることがわかる。

太原氏等が大学での英語教育の課題を研究テーマとしたことを高く評価したいと思う。英検の研究助成では、小・中・高の英語教育を研究対象とするものが、圧倒的に多く、また、研究助成の報告書を読む英語教育関係者は小・中・高の英語教育関係者が中心である。しかし、日本の学校英語教育の最終段階が大学における英語教育である以上、英検の研究助成に大学英語教育の研究が欠かせないと思う。

特に、大学での英語教育と接点が強い高等学校における英語教育関係者にとって本論文は参考になるところが多い。まずは、研究をする手続きが詳細である点が参考になるであろう。先行研究への目配り、先行研究の位置付けと評価、データの処理など参考になるところは大である。一方、高等学校の英語教育関係者は、本論文から大学のどの分野の学生や卒業生を研究対象としたのか、その結果はどうだったか、などの情報が知りたいのではないかと思う。研究対象が「英語教育・英語学・英米文学」の学部・大学院の学生が主体のようだが、ほかの学部や大学院の場合の情報も知りたいのではないかだろうか。今後のさらなる研究を期待する。

C. 調査部門・報告Ⅱ

吉田 研作

日本人大学生の英語学習における
自己調整学習能力尺度の開発：
英語資格試験に向けた自主学習に
焦点を当てて

【研究者：福田 晶子】

何かを学ぶということは、それを学ぶための情意的要因、動機づけに関する要因、学習方略にかかる要因が必要で、それらが有機的に働くことで、自己調整学習能力が効力を発する、と解釈される。つまり、本稿では、自己調整学習を「学習者が自らの認知・情意・行動・動機をコントロールし、維持しながら学習を進める過程」(Schunk& Greene)と捉え、そのための有効な能力測定方法を開発することを目指している。なお、本稿では、学校や教室での勉強ではなく、自らが何かを学ぶ際に適応されるもので、特に外部テスト(ここでは TOEIC)を受験する際の状況が設定されている。

研究の結果を見ると、従来、楽しさや知的好奇心、達成感などの学習者の内的要因は学習にポジティブに影響するが、例えば、他者に対する優位性、テストの結果などの外的評価などの外的な要因はどちらかというとネガティブに影響するといわれてきたが、外部テストを受験するという設定の下では、外的要因も有意に働くことが示されている。つまり、自主学習においては、従来考えられていた以上に様々な要因がかかわっているということが分かったという。

本稿の特徴は、自己調整学習能力尺度の開発にあるが、そのために非常に綿密な研究及び分析手法を用いている点にある。被験者が大学生である点、また、TOEICという一つのテストのための受験準備という限定はあるものの、研究手法、統計的分析方法の適用という観点から大いに参考になる研究だと言えるだろう。

第32回 研究助成

A 研究部門・報告 I・英語能力テストに関する研究

ライティングタスク(技能独立型 vs. 統合型)が 発表語彙とその測定に与える影響: TAALESによる語彙の洗練性分析を基に

研究者:茨城県／筑波大学大学院 在籍 小室 竜也

《研究助言者:寺内 一》

概要

本研究は英作文に含まれる発表語彙の洗練性を数値化するツール Tool for Automatic Analysis of Lexical Sophistication(TAALES)を用いて、技能独立型および技能統合型のライティングタスクと発表語彙のかかわりを検証した。

調査1では英検1級から3級で用いられているテキストを対象とし、受験級ごとにどのような洗練性指標が分類を予測するかを分析した。その結果、上位の受験級では心理言語学的な要因である長い反応時間や英語母語話者の習得年齢が高い語彙が使用されていることがわかった。

調査2では大学生を対象に独立型および統合型のライティングタスクを実施し、タスクの分類を予測する洗練性指標を分析した。その結果、発表語彙における語と語の結びつきの強度、音韻的に類似する語彙、内容語の多義性の3点がタスクを分類するのに寄与していることが示された。

1

研究の背景

4技能5領域の指導の中でも、リスニングとリーディングは受容技能、スピーキング(発表・やりとり)とライティングが産出・発表技能と分類される。さらに、リーディングが語彙認知や統語解析、推論生成などに下位区分できるように、語彙知識も連語や形態素などのいくつかの側面を持っていると考えられている(Nation, 2013)。

ライティングやスピーキングにおいて必要とされる語彙知識は発表に関する側面であり、発表語彙知識はスピーキングおよびライティングテストの得点を説明する重要な要素である(Kyle & Crossley, 2015, 2016)。ライティングにおける発表語彙を測定する方法の1つに、実際に生徒が書いた英文をいくつかの観点から分析するというものがある(Laufer & Nation, 1995)。この方法で多く用いられるのは特定のトピックが与えられ、それについて自分の意見を自由に書く技能独立型の形式である。現在、日本の英語教育の実践現場では4技能の統合的な指導が行われているが、複数技能を組み合わせたライティング活動における発表語彙の特徴は十分に明らかではない。そこで本研究では技能独立型と技能統合型のタスクの違いによって、作文中で用いられている発表語彙はどのような影響を受けているのかを検証することとする。

本稿では2つの調査結果を報告する。調査1では、特定のトピックについて自分の意見を述べる形式(技能独立型)である英検のモデル英作文を語彙の洗練性分析ツール(Tool for Automatic Analysis of Lexical Sophistication; TAALES)によって調査することにより、各受験級における発表語彙の特徴を把握する。続く調査2では技能独立型および技能統合型(トピックに関する英文を読解後、自分の意見を書く)タスクを実施し、タスクの違いによる影響が発表語彙の特徴に与える影響をTAALESによって調査する。これらの2つの調査を通し、4技能を総合的・統合的に指

導する教育実践者に対して示唆を与えることを目的とする。

2 先行研究

2.1 発表語彙の測定

新学習指導要領では高校卒業時までに、小学校からの合計で約4,000から5,000語の語彙を指導することが示されている。語彙知識が言語運用能力の基礎の1つであることは多くの教育実践者に受け入れられており、学習者自身も語彙知識の重要性を自覚している。その語彙知識の中でも、低頻度の難しい語彙の知識を有していることは熟達した学習者であることを意味しており、洗練された語彙を知っている・使うことができるることは言語熟達度の指標の1つとして機能する(Lu, 2010)。この考えに基づき、学習者がどのくらい語彙をたくさん知っているか測定する診断テストが多く開発された(e.g., Vocabulary Size Test; Nation & Beglar, 2007; New Vocabulary Level Test: Webb, Sasao, Ballance, 2017)。

日本人英語学習者の語彙サイズを診断するテストとして最も一般的なのは望月語彙サイズテスト(相澤・望月, 2010)である(図1)。これは頻度レベルごとに日本人英語学習者が知っておくべき8,000語の語彙をリスト化した大学英語教育学会基本語リストJACET8,000(石川他, 2003)に基づき、1,000語レベルの頻度バンドごとにランダムに抽出した語彙の意味を選択肢の中から選ぶ。例えばcraftsmanは5,000語レベルに属しているため、craftsmanの意味を知っている学習者は5,000語よりも簡単な頻度の高い語彙は知っていると見なすことができる。しかし実際の測定は複雑であり、3,000語レベルの問題で誤りがあった学習者が7,000語レベルの問題をたまたま解答することができる可能性もある。このようにいくつかの問

題点は抱えているが、受容技能に関する語彙知識を測定する診断テストは様々なものが開発されている。しかし、学習者の発表語彙を測定する方法は受容語彙に比べてあまり発達していない。

発表語彙の測定には大きく2種類のアプローチがある。1つはこれまでに開発されてきた受容語彙サイズテストを援用し、発表語彙知識を直接測定するものである。もう1つは作文やスピーチといった学習者がアウトプット活動で実際に産出した成果物を間接的に分析する方法である。本節ではそれぞれの利点と欠点について総括する。

2.1.1 直接的な測定

Laufer and Nation(1999)のProductive Level Test(PLT)では、受容語彙サイズを測定するVocabulary Level Test(Nation, 1983)を援用した発表語彙サイズテストである。このテストでは頻度レベルごとに目標語がランダムに抽出され、適切な目標語だけが正しく産出されるように目標単語の一部の文字を提示した上で、受験者は目標語を母語から第二言語へ翻訳する。例えば、Laws are based upon the principle of jus_____の中で目標単語 justice の数文字が提示される。この方法は頻度に基づいて、発表語彙知識を直接測定していると言える。数文字を与えることにより、テスト作成者は厳密に頻度レベルを統制することができるため、具体的な数値で発表語彙サイズが算出され直感的に理解しやすい。さらに直接的に発表語彙知識を測定することにより、受験者が自信のない語彙の使用を回避することなくすことができる。

PLTは効率よく発表語彙知識を測定できるが、以下の2点の問題点がある。まず実際には外部から刺激が与えられた状態でのみ使用できる語彙知識を測定しているため、受容語彙知識の単なる翻訳テスト版であるとも考えられる。次に、その結果が必ずしも言語使用の場面で目標語が使えるかどうかとは関係がない可能性である。学習者が自らの意思で使いこなすことができる発表

■図1: 望月語彙サイズテストの例(相澤・望月, 2010)

1. 持っている	2. しなければならない
(1) do (2) get (3) give (4) have (5) must (6) raise	

語彙知識を測定するためには、実際に英作文を書かせたりスピーチをさせたりしたアウトプットを分析する方法が必要である。

2.1.2 間接的な測定

発表語彙知識を測定する2つめの方法としては、学習者のアウトプットを様々な観点から分析する間接的測定がある。具体的には、英作文で使用されている語彙を4つの頻度レベル(1,000語, 2,000語、大学語彙、それより頻度の低い語)に分け、それぞれのレベルに含まれる語彙の割合を表す語彙頻度プロフィール(Laufer & Nation, 1995)が多く用いられている。例えば、英作文100語のうち70語が1,000語レベル、20語が2,000語、5語が大学語彙、5語がその他の固有名詞で構成されるとすると、70%-20%-5%-5%と表される。

語彙頻度プロフィールを発表語彙能力の測定方法として考える前提には、熟達した学習者は難しい語彙を知っており使いこなせるというものがある。つまり、I think… I think…のように高頻度語を繰り返し使う学習者よりも、I suppose… I assume… のように低頻度語を使うことができる学習者の方が熟達した発表語彙知識を有していると言える。なお、語彙頻度プロフィールに必要となるプログラムは Paul Nation 氏のウェブページにて無料で使用可能である。他にも P_Lex というツールも無料で使うことができる(図2)。

語彙頻度プロフィールや P_Lex は言語使用的な場面において産出された語彙を分析することか

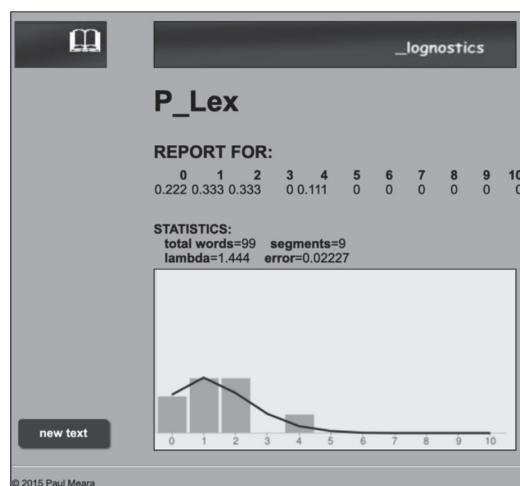
ら妥当性が高い一方、発表語彙サイズが大きいのかどうかが直感的に分かりにくい。このデメリットを改善したのが Kojima and Yamashita (2014) の S (http://kojima-vlab.org/lexical_richness/S_English.html) というプログラムである(図3)。S は最終的なアウトプットで使用されている語彙を頻度レベルに基づいて分析し、書き手の発表語彙知識を算出する。例えば、ある生徒が書いた作文を S のプログラムで分析すると、2795.46 という数値が出てきたとする。このことが意味するのは、書き手の発表語彙サイズは 2795.46 語であるということである。しかし、これらのプログラムは頻度レベルという観点のみに基づいており、語彙の豊かさの観点からは問題がある。

2.2 語彙の豊かさ (Lexical Richness)

語彙の豊かさという用語は1人の書き手が使用する語彙の数のことを指している。近年の第二言語でのライティングやスピーキングの研究の評価指標には複雑性、正確性、流暢性に語彙の豊かさという概念を加えた指標が用いられることが多い(Polio & Friedman, 2016)。

2.2.1 語彙の数え方

I like playing tennis. I am a member of the tennis club. I play tennis every day. という英作文ではすべての語を1語と数える方法で17語となる。しかし、人称代名詞の I と普通名詞の tennis は



■図2: P_Lexの分析結果の画面

```
S
2795.461
Formula: HIRITU ~ (log(FREQ)/log(S)) * 100
Parameters:
Estimate Std. Error t value Pr(>|t|)
S 2795.46 38.45 72.7 <2e-16 ***
-
Signif. codes: 0 **** 0.001 *** 0.01 ** 0.05 * 0.1 ' ' 1
Residual standard error: 3.496 on 494 degrees of freedom
Number of iterations to convergence: 5
Achieved convergence tolerance: 2.176e-10
```

■図3: Sの分析結果の画面

3回登場している。分かち書きされた語を1語とカウントする方法は総語数(token), 重複を1語とする数え方は異語数(type)とそれぞれ呼ばれている。

2.2.2 語彙の多様性

多様な種類の語彙を使うことができる学習者は、熟達した発表語彙知識があると考えられる(Laufer & Nation, 1995)。例えば、上述のI thinkだけでなく I suppose や I assumeなどの様々な表現方法を使いこなすことが熟達した書き手には可能である。このことから、語彙の多様性が高い英作文は高く評価されると言える。この多様性の指標には伝統的に Type を Token で割った Type Token Ratio(TTR) が用いられることが多い。

しかし、TTRには語数が長くなれば多様性は減少してしまうという問題点がある。例えば、熟達度の高い学習者は流暢性が高いため多くの語彙を産出することができるが、多くの語彙を産出すると冠詞や前置詞の割合が高まり、多様性が低くなってしまう。つまり、本来であれば熟達度が低いと評価されるはずの、流暢性が低い作文やスピーチをした学習者の方が多様性のある語彙を用いることができると評価されてしまう。文章の長さに影響を受けてしまうという TTR の欠点を改善するために Guiraud Index や Measure of Textual Lexical Diversity(MTLD), vocD, HD-D などの様々な指標が提案され(Kyle, 2019), ライティング研究ではその特徴に応じてそれぞれの指標が報告されている。

語彙の多様性は、作文中で使用されている語彙の難易度が低い場合の判断に根本的な問題がある。例えば、I love tennis very much. I want to be like Nishikori. という 10 type 11 token の作文は $10/11 = 0.91$ TTR と計算され、The rate of juvenile delinquency is increasing at an alarming rate. という 10 type 11 token の作文と同じ多様性レベルであるとみなされてしまう。この問題点の解決策に語彙の洗練性指標が挙げられる。

2.2.3 語彙の洗練性

洗練された語彙は高度で難しい単語であると一般的に考えられ(Lafuer & Nation, 1995), 語

彙の多様性と類似した傾向を示すが、実際には語彙の洗練性は多様性とは異なるものである(Jarvis, 2013)。例えば、1人は3,000語レベルの易しい語彙を使って100語の作文をし、もう1人は5,000語レベルの難しい語彙で100語の作文をしたとする。前者は上限が3,000語レベルであるため5,000語レベルの語彙を使う後者よりも語彙の多様性は低くなり、高頻度の易しい語彙を使っている前者の方が洗練性も低いと言える。しかし上述の The rate of juvenile delinquency is increasing at an alarming rate. の例のように、作文が同程度の多様性を示す場合には洗練性の指標を参照する必要がある。つまり、語彙の多様性が作文全体に占める特定の語彙の割合であるとするならば、語彙の洗練性は作文全体に占める難易度の高い洗練された語彙のことを指す(Kim, Crossley, & Kyle, 2018)。

語彙の洗練性は多面的であり、難易度の高低は簡単に決めることはできないが、歴史的には頻度に基づいて研究が発展してきた(e.g., Laufer & Nation, 1995)。2.1.2にて触れたように、作文中で5,000語レベルの語彙の割合が高くなれば洗練された語彙を使用しており、その書き手は熟達した発表語彙知識を有していると評価できる。

しかし、語彙の洗練性は頻度という観点のみでは定義することが難しい(Kyle & Crossley, 2015)。近年の研究では、頻度以外の観点として大きく次の4つが挙げられる(Kyle & Crossley, 2016)：①range(母語話者の参照コーパスにおいて、その語彙が収録されているファイルの数；McNamara, Crossley, & McCarthy, 2010), ②n-gram 頻度(生起する n 語の連続の頻度；Crossley et al., 2012), ③学術語彙(Coxhead [2000] や Simpson-Vlach and Ellis [2010] の学術語彙リストに基づく頻度), ④心理言語学的要素(具象性、親密度、心象性、有意性、Age of Acquisition; Crossley & Skalicky, 2019)。他にも、階層性(上位語、下位語)や多義性、正書法・音韻的近傍語などの観点もある。

語彙の洗練性は多面的であるが、これらは類似した構成概念を測定している可能性がある。例えば、低頻度の語彙は親密度が低く、それゆえ母語話者であっても反応時間が長くなる。反応時間が長い語彙は綴りと音の対応関係が不規則で

あり、スペルが似たような単語である近傍語が多く存在する。さらに「ワイン」や「チーズ」のように日本語化された英単語も多く存在し、英単語としては頻度が低くとも日本人英語学習者はカタカナで知っている場合もある。それゆえ、語彙の洗練性に関する明確な定義は未だ議論が続いている。

2.3 技能統合型ライティングタスク

語彙知識は母語や第二言語のライティング能力において重要な役割を果たしている(Baba, 2009; Kim & Crossley, 2018)。ライティング能力は言語知識と方略的能力に大別され(Weigle, 2002), 発表語彙は前者の下位要素であり、特に洗練性は言語知識の重要な構成概念である(Bachman & Palmer, 1996)。なお、作文の目標設定やトピックに関する背景知識などは後者に分類されている。実際に、高等学校学習指導要領では文法的正確性や語句の選択だけでなく、書く前のブレーン・ストーミングや書いた作品を読み返して推敲するなどの養成も視野に入れている。つまり、作文をする学習者は使用する語彙の洗練性や文法の正確性だけでなく、書く題材に関する背景知識などを十分に活性化するなど、ミクロだけでなくマクロな視点を持つことが求められている。さらに第二言語での作文の場合、言語知識に認知資源を多く配分してしまうため母語で作文するときよりも内容に対してはあまり注意を向けることができない。英作文の指導者や評価者はこれらの点を経験的に理解しているが、明示的に認識することが重要である。

上述のライティング能力の構成概念は技能独立型の活動において当てはまるものであり、技能統合型とは異なる。技能独立型のライティングは例えば、「高校で制服を着用することに賛成か反対か」というトピックのみが与えられ、書き手自身でアイディアを生成しなければならないものである。しかし教室外の英語使用の場面に目を向けると、単に自分の考えや意見を書くだけでなく、Eメールの返信を書いたり、読んだ本の感想を書いたりする。このような技能統合型の活動は、独立型と比べて書き手に要求する能力が異なっている。例えば、独立型では提示された絵に

ついて感想を書く一方、統合型では提示された絵とその絵の作者が込めたエピソードなどの文章を読解(聴解)し、その内容に基づいて作文する。統合型ではリーディング・リスニングにおいて単に理解するためだけではなく、作文することを念頭に置いて情報を選択するため受容技能的重要性が高い(佐藤, 2017)。ただし、技能統合型の活動で必要となる能力は単なる受容技能と発表技能の組み合わせではなく、選択した情報に基づいた発表が求められている(Plakans, 2013)。

複数の技能を統合したとしても、評価の対象となるのは最終目標を設定した最後の技能であるライティングとなる(望月, 2015)。この理由は、生徒一人ひとりのリーディングの内容理解に加え、作文の評価をすることは現実的・実用的ではないためである。例えば、TED-Edの動画の内容をメモしながら最終的に要約をするという活動では、学習者が内容を理解したかどうかを確認する方法は学習者のとったメモとなる。しかし、5分の動画であっても30名のクラス全員のメモとアウトプットの両方を採点するのには時間が非常にかかる。それゆえ、最終的なアウトプットを評価することが妥当性・実現可能性の観点から一般的である。

2.3.1 技能統合型の問題点

技能統合タスクを教室内で実施する際には、技能間の統合の強さに着目することが有用である(根岸, 2011)。例えば、Eメールを返信する場合は受信したメール内容を理解する必要性があるため、技能が強く結びついていると考えられる。一方、運転免許取得の年齢を引き下げることについて賛成意見を読み、それを参考に自分の意見を書くタスクの技能統合は弱いと言える。

技能統合型タスクは英語学習者が将来に出会う可能性の高い言語使用場面を想定しているため真正性が高く、教室内で多く取り入れられると考えられる。しかし、2つの欠点がある(Plakans, 2013)。1つはある程度高い熟達度が必要となることである。技能の統合が強い場合、提示されているソースを十分に理解できなければ何もアウトプットすることができない。つまり受容技能の熟達度が低く、例えばリスニングで提示されるキーワードを聴き逃してしまった場

合にはスピーチング能力がある程度あったとしても産出する内容がないため、熟達度を低く評価されてしまう。幅広い熟達度の受験者がいる英検で技能統合型テストが実施されていない理由の1つに、このデメリットが関係していると言える。もう1つのデメリットは発表技能そのものの診断目的には向かないというものである。例えば、生徒が技能統合型タスクで十分に作文することができなかった場合、その原因はライティングに関わる発表語彙知識が不十分だったためなのか、リーディングに関わる受容語彙知識が足りなかつたためなのかが曖昧となってしまう。

一方で技能の統合が弱い場合、情報源となる読解テキストが存在することで学習者の文章構造や内容に関する処理を手助けし(Plakans, 2008)、不慣れなトピックに関する情報資源を書き手に提供することから(Gebril, 2009)、独立型よりも多くのアイディアを生み出し自らの主張を裏付けることができる。技能統合型ライティングに関する研究からは、語彙を始めとする言語的なサポートを得ることができるするために作文中にソーステキストに読み戻ることが多いことや(Gebril & Plakans, 2016; Plakans & Gebril, 2012)、文法の複雑性(Biber et al., 2016)および語彙の洗練性が高く(Kyle & Crossley, 2016)、結果性の高さがライティングの得点を予測する(Guo, Crossley, & McNamara, 2013)といった特徴が明らかになっている。以下では発表語彙の洗練性に焦点を絞って統合が弱い技能統合型ライティング活動と語彙の洗練性の関係について明らかになっている特徴を述べる。

2.3.2 ライティングタスクと語彙の洗練性

第二言語学習者の英作文において、洗練性の高い難しい語彙を使ったものはライティング熟達度を判断する指標の1つとして機能する(Laufer & Nation, 1995)。この考えには、簡単な英語であっても作文の内容を重視するべきであるという批判があるかもしれないが、学術目的の英語や医療や司法といった特定の目的のための英語を用いる場合には洗練性の高い作文能力が求められる。このような背景から、近年では語彙の洗練性を分析する TAALES を用いた研究が注目されている(Crossley et al., 2019; Kim et al.,

2018)。TAALES は Web 上のサービスである Coh-Metrix と異なり、使用者の PC にダウンロードする自然言語処理ツールの1つであり、400 を超える大量な洗練性指標を高速で提示する。

2.2.3 で述べたように語彙の洗練性は多面的であり明確に定義することは難しい。技能独立型ライティングで洗練性が重要な指標であることが述べられていても、それは頻度に基づくものなのか、それとも反応速度のような心理言語学的な要素に基づくもののかは明らかにされてこなかった。そこで Kyle and Crossley (2016) は複数の観点から語彙の洗練性を算出するため TALES を用いて、技能独立型と技能統合型ライティングタスクを比較した。その結果、技能独立型テストにおいては、得点の分散の 36.8% は語彙の洗練性(BNC Written Range で 16.7%, BNC Written Bigram Frequency Logarithm で 17.8%)によって説明された。このことは、言語知識の下位区分である語彙の洗練性の重要性を支持するものである。説明率の多くを占めていた Range は、低いほど特定の領域において使用される専門用語であることを意味しており、高い洗練性とみなされる。bigram は2語の連語のことを指し、適切なコロケーションを使用できているかどうかを意味している。連語の知識は語彙のネットワークだけでなく、文法知識とも結びついていることから(Sinclair, 1991), bigram 頻度が高い方が語彙的にも文法的にも豊富な言語知識を有しているとみなされる。

一方で、技能統合型ライティングテストは独立型よりも高い洗練性を示すが、得点の分散の 8.3% しか語彙の洗練性によって説明されなかった(Kyle & Crossley, 2016)。この理由について、Kyle and Crossley は Baba (2009) の結果について言及している。Baba (2009) では受容語彙知識の量と質が技能統合型ライティングの一種である要約課題において中程度の相関があること ($r = .400$, $r = .340$) を報告している。Plakans (2013) で示されている通り、技能統合型の活動は単なる「受容 + 発表」というわけではないが、提示されているソーステキストを理解するために受容技能は一定の熟達度が必要となる。さらに、ソーステキストからの言語的な借用が可能であるため、文法的な誤りが少なく正確性が高いこ

とや、アイディアの生成を助けていたため流暢性が高いことも得点には関係していると考えられる。

2.4 Tool for Automatic Analysis of Lexical Sophistication

語彙の洗練性を加味して発表語彙を自動で算出するツールには語彙頻度プロフィールやP_Lexなどがあるが、これらはいずれも頻度の指標のみに基づいている。洗練性が多面的であることを踏まえると、range や bigram、心理言語学的な指標などを含めて詳細に分析する必要がある。TAALES はこれらの観点を含め、version 2.2 では484の指標(2020年5月現在、version 2.8.1のベータ版で1100を超える指標を算出可能)を提供するツールである。このツールは、①誰でも無料で使用することができ、②Windows, macOS, Linuxといった複数のOS上で動作させることができる。TAALES のインストールは NLP for the Social Sciences のウェブサイトから行うことができる。

TAALES で使用されている指標は次のとおりである。

(1) 頻度

語彙の洗練性指標として歴史的に用いられてきたのは、コーパスにおける頻度である(e.g., Laufer & Nation, 1995)。TAALES が参照するコーパスは BNC, SUBTREXus, COCA, MRC Psycholinguistics Database である。British National Corpus(BNC)は約1億語のイギリス英語を収集した均衡コーパスであり、SUBTREXus は映画字幕のコーパス(話し言葉)である。Corpus of Contemporary American English(COCA)は2020年5月現在、8つのジャンル(spoken, fiction, popular magazines, newspapers, academic texts, TV and Movies subtitle, blogs, web page)からサブコーパスが構成されている約10億語の現代アメリカ英語のコーパスである。なお、TAALES(version 2.2)では academic, fiction, magazine, news, spoken の5つから選択する。最後の MRC Psycholinguistics Database は心理言語学で用いられる指標だけでなく(e.g., familiarity, meaningfulness, imageability),

シラブル数やストレスパターンなどについても26のオプションから細かく条件が設定されている。

(2) Range

Range はコーパスに収録されているテキストの数のことを指す(Gries, 2008)。例えば、read や make といった語彙はどのようなテキストでも出現する可能性が高いため、range は高い。一方で、extremist や adherence といった語彙は出現するテキストの範囲が限定的であると考えられる。それゆえ、range の値が低い語彙は洗練性が高いと考えられる。

(3) N-gram 頻度

N-gram とは一定の文字数で構成される語の塊のことであり、n は任意の数字を示す。例えば、bigram は2語の塊を意味している。kick the bucket のようなイディオム、on the other hand や in my opinion といった連語表現を自由に使うことができる学習者は熟達度が高いと判断することができる(Syanova-Chantria & Pellicer-Sánchez, 2018)。N-gram のような連語に関する知識は単語同士のつながりを表し、語彙知識の質を指す。N-gram を含む語と語の結びつきの強度は、MIスコア(mutual information score、相互情報量)という統計値によって計算される。MIスコアは中心語と共起語が互いに相手の情報をどの程度持っているかを表す。つまり「Xと言えばY」「Yと言えばX」のように、片方の語それぞれに、あらかじめ互いの語の情報を取り込まれているような強い結びつきが成立していることが評価される。

(4) 学術語彙

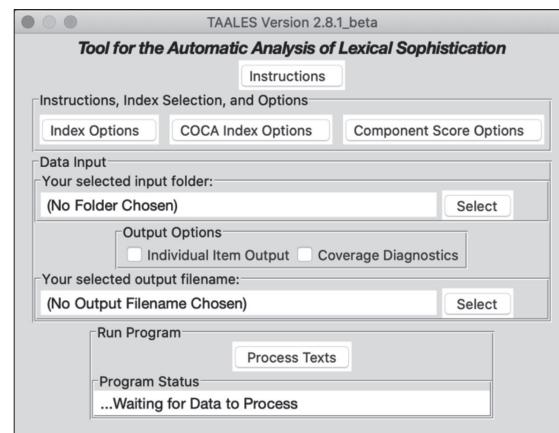
学術場面において使用される語彙は難易度が高く、洗練性が高い。学術語彙をまとめた語彙リストとしては Coxhead (2000) の Academic Word List や Simpson-Vlach and Ellis (2010) の Academic Formulas List が挙げられ、TAALES の参照コーパスとして用いられている。

(5) 心理言語学的要素

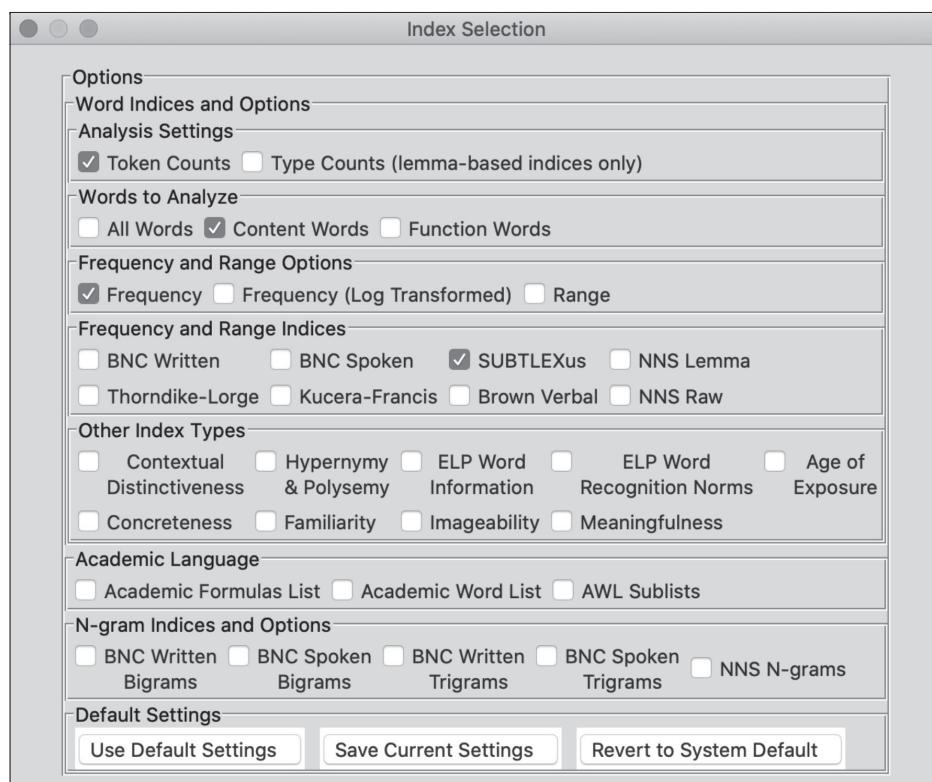
第二言語学習者の語彙知識は長期記憶に格納されており、この語彙に関する記憶体系はメン

タルレキシコンと呼ばれている。このメンタルレキシコンをモデル化する心理言語学の手法として、提示された目標語が非単語か実在語かを判断する反応時間を測定する方法が一般的に用いられている。反応時間に関する研究からは、目標語の特徴によって①頻度効果(反応時間は高頻度語よりも低頻度語の方が長い)、②正書法準拠効果(正書法に準拠した非単語はランダムな文字列の非単語よりも反応時間が長い)、③近傍語サイズ効果などの影響が得られることが明らかになっている。近傍語とは単語を構成する1文字を同一文字位置で別の文字に置き換えた際に得られる単語のことを指す(Andrews, 1989)。例えば、bank の近傍語は rank, tank, sank などである。近傍語の数が多いほど反応時間は長くなる。TAALES では MRC Psycholinguistics Databaseに基づいてこれらの心理言語学的指標を算出する。Guo et al.(2013) は心理言語学的要素を Coh-Metrix で算出し、第二言語学習者の技能統合型英作文の得点の分散の8.5%を説明することができたことを述べている。

以上のように、TAALES を用いることにより様々な観点から語彙の洗練性を分析することができる。図4および図5に TAALES の使用画面を示す。.txt ファイルを用意すればマウス操作で一括で簡単に分析することが可能である。



■図4: TAALES version 2.8.1 (ベータ版) のインターフェイス



■図5: TAALES version 2.8.1 (ベータ版) の指標選択画面

2.5 本研究の目的と構成

発表語彙知識の測定には、語彙 サイズテストを用いる直接的な方法と英作文を用いる間接的な方法がある。前者は実際の言語使用を反映していない可能性がある点、後者は頻度の観点のみに基づいている点にそれぞれ問題がある。さらに、近年の技能統合型活動に注目が集まっている中で、技能統合型ライティングタスクと発表語彙との関係は不透明であった。そこで本研究では、発表語彙の間接的なアプローチの問題点をTAALESによって算出される豊富な洗練性指標を用いることによって解決し、技能独立型と統合型のタスクの違いが発表語彙に与える影響を明らかにする。この目的を達成するため、本研究では2つの調査を行う。調査1では技能独立型である英検のモデル英作文における発表語彙の特徴を明らかにする。調査2では、技能独立型および統合型タスクが発表語彙に与える影響を調査する。

3 調査1: 英検のモデル英作文の分析

調査1では英検のモデル作文における発表語彙の特徴をTAALESによって数値化し、各受験級で求められる語彙の洗練性がどのように異なるのかを明らかにすることを目的とする。そのため1級から3級までのモデル英作文を用いて、発表語彙に関わる洗練性指標はどのように異なるかを検証する。検証課題(Research Questions: RQs)は次の2点である。

RQ1-1 TAALESによって算出される洗練性指標のうち、どの指標が受験級を予測するか。

RQ1-2 TAALESによって算出される洗練性指標は受験級によってどのように異なるか。

■表1: 調査1で分析したモデル英作文の特徴

受験級	n	平均語数		FKGL		S	
		M	SD	M	SD	M	M
1級	22	226.59	16.59	12.49	1.24	11684.18	2069.93
準1級	13	147.38	2.56	11.07	1.00	5319.24	2518.29
2級	13	92.77	4.02	7.17	0.88	2014.4	632.57
準2級	10	57.1	2.7	4.58	1.24	1870.09	602.06
3級	10	31.5	1.96	4.8	1.60	-	-

(注) S(Kojima & Yamashita, 2014)による分析で妥当な結果を得るために最低50語が必要であるため、3級の結果は記載していない。

■表2: 調査1の記述統計量

受験級	1級	準1級	2級	準2級	3級
洗練性指標	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
LD Mean RT	646.90 (8.56)	629.66 (3.92)	614.33 (4.85)	609.28 (6.09)	601.20 (7.81)
Kuperman AoA CW	7.39 (5.66)	6.69 (4.48)	5.60 (3.00)	5.03 (3.93)	4.72 (4.31)
MRC Familiarity CW	569.74 (3.85)	576.22 (5.76)	592.64 (5.43)	596.24 (8.09)	594.84 (10.23)

3.1 方法

3.1.1 英検テキストの収集

TAALESによるテキスト分析を行うため、英検1級から3級までのモデル英作文を収集した。分析対象としたテキストの詳細および、頻度の観点に基づくSの結果を表1にまとめる。この結果は受験級が高くなれば使用されている語彙も低頻度語彙であることを示している。

3.1.2 手順と分析

TAALESによるテキストの分析に先立ち、固有名詞や記号などは除去された。その後、洗練性のどの指標が英検の受験級を予測するのかを知るために語数を除いた指標で決定木分析が行われた。決定木分析は説明変数を用いて段階的にグループ分けをしていく分析方法である。今回の説明変数は TAALESによって算出される洗練性の指標である。決定木のノード(枝分かれ節点)に含まれなかった指標について詳細に検証するために多変量分散分析(multivariate analysis of variance: MANOVA)を行った。MANOVAは複数の独立変数と2つ以上の従属変数を分析することができる手法である。なお、TALESによって算出される洗練性指標は484という多量の指標が示されるため、決定木のノードおよびMANOVAの従属変数に含まれた指標以外の記述統計については割愛する(表2)。

3.2 結果と考察

決定木分析の結果を図6に示す。英検の級の分類を最も予測したのは Response Time(RT: その語が実在語か非単語か判断するために有する反応時間のことを指し、反応時間が長いものは洗練された難易度の高い語彙であると考えられる)であった。英検1級では635.883ミリ秒以上の反応時間を有するような洗練された語彙を使うことができることが求められる。第2および第3のノードは Age of Acquisition(AoA: その語を母語話者が習得する年齢のことを指し、高い年齢で習得する語彙ほど洗練された語彙であると考えられる)であった。つまり英検準1級では5.417歳以降、英検2級では4.835-5.417歳に習得する語彙を使える

ことが求められているという結果が得られた。

これらのノードはいずれも心理言語学の指標であり、母語話者を対象にしたデータベースを基にしている。RTは語彙認知やメンタルレキシコンへのアクセススピードにも関連する指標、AoAはインプットの頻度と関連する指標である。前者の結果から、母語話者であってもその単語が存在するかどうか6秒以上判断に迷う単語を英検1級のモデル作文で使用しているため、要求される発表語彙のレベルは非常に高いと言える。また、アクセススピードには頻度効果が影響する。第2ノードとなっている AoA の指標は、母語話者が生まれてから5歳ごろまでに触れるインプット量とほぼ同義であり、そのインプットには頻度の高いものと低いものが含まれている。それゆえ、第1ノードと第2ノードは頻度という共通点を持つ類似した構成概念であると考えられる。

認知言語学で提唱された用法基盤モデルに基づくと、外部から提供されるインプットの頻度が高いものほど習得される傾向にある。この頻度効果に関して、Conklin(2019)は低頻度語彙ほど頻度効果の恩恵が大きいことを述べている。例えば、1回の授業の中で高頻度のreadやgiveを繰り返し登場させるよりも、低頻度語彙であるextremistやadherenceを複数回登場させる方が遭遇回数の効果が大きく見える。しかし、授業時間が限られている日本の教室現場の多くでは、頻度効果が見られるほど大量のインプットは提供されていないのが現実である。むしろ語彙は明示的な方法で指導されることが多い。さらに、明示的・意図的に指導される語彙は言語使用場面を想定し、高頻度で用いられるものであることが多い。

今回の決定木分析の結果は、洗練性指標が歴史的に頻度に基づくものであったこと(e.g., Laufer & Nation, 1995)とも一致する。この分析で明らかになった点は、頻度指標が単なるコーパスでの出現頻度ではなく、英語母語話者の反応時間や習得年齢と関係することである。つまり、モデル作文の作者が母語話者であった可能性や言語使用場面で高頻度な語彙は母語話者も使用することが関係している可能性も考えられる。RQ1に対して、受験級の予測には反応時間と習得年齢の2つが寄与しているという解答が得られた。

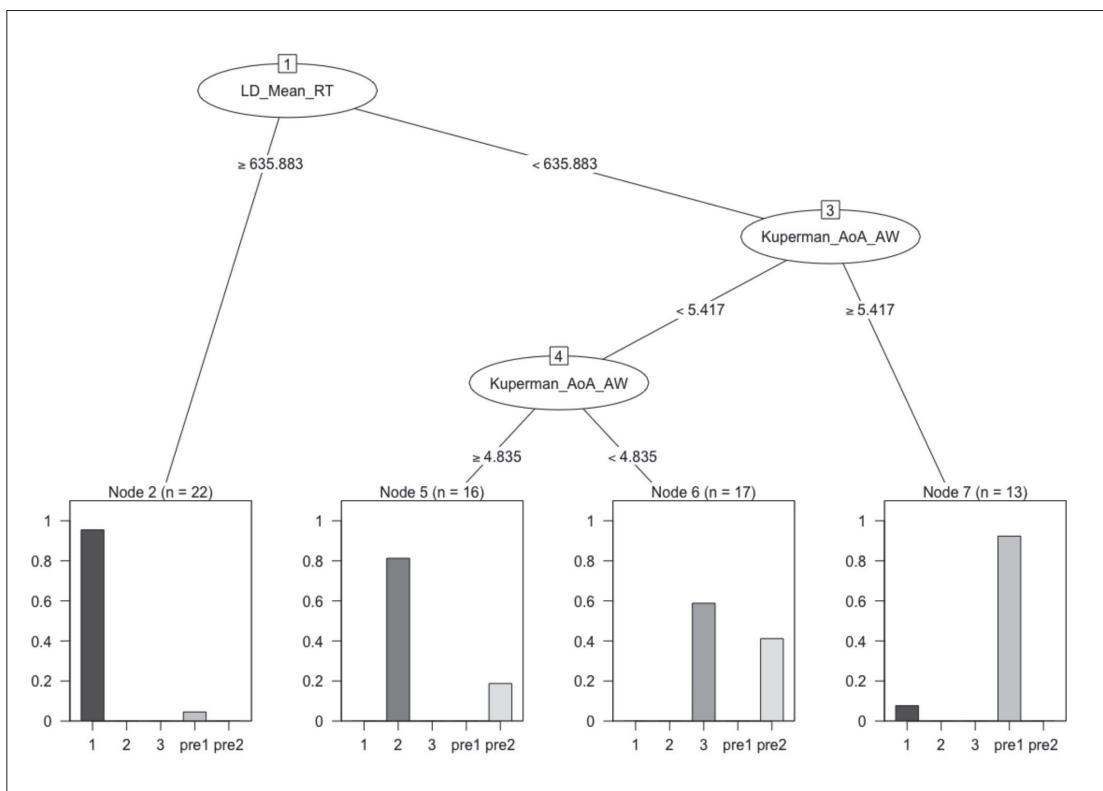


図6: 調査1の決定木分析の結果

RQ2は受験級の間で洗練性指標はどうに異なるのかを扱っている。このRQに解答するため、対応のない1要因の計画で、決定木のノードを独立変数としてMANOVAを実施した。なお、英検1級のみサンプル数が多かったことから、BoxのM検定において異質と判断されないようにサンプルをランダムに削除し、他の受験級とサンプル数を等しくした。記述統計を表2に示す。Pillaiのトレースによる結果に基づくと受験級の洗練性指標は有意に異なっていた、 $F(12, 135) = 7.43, p < .001, \eta^2 = .398$ 。共分散行列の等質性は非有意であったが、Leneneの検定でMRC Familiarity CW(内用語の親密度)が5%水準で有意であり、等分散性が満たされていなかった。それゆえ、結果の解釈は決定木分析のノードに現れた2つに限定する。1変量分散分析による各指標の結果、LD Mean RTで $F(4, 45) = 78.36, p < .001, \eta^2 = .87$ 、Kuperman AoA CWで $F(4, 45) = 64.62, p < .001, \eta^2 = .85$ は0.1%水準で有意であった。それぞれの多重比較を見ると、反応時間は1級と準1級($p < .001$)、準1級と2級($p < .001$)で高い受験級で有意に反応時間の長い語彙を使用しており、2級と準2級($p = .558$)、準2級と3級($p = .120$)では有意な差が見られなかった。この傾向は習得年齢についても同様であり、1級と準1級($p = .024$)、準1級と2級($p < .001$)でそれぞれ高い受験級で習得年齢の高い語彙が使用されていたが、2級と準2級($p = .101$)、3級と準2級($p = .648$)では非有意であった。この結果と図6を組み合わせて考えると、母語話者が5歳前後で受容語彙として習得する語彙は英検1級の発表語彙として求められており、大人の母語話者であっても6秒以上判断に迷う語彙は英検1級の発表語彙の特徴であると言える。しかし、これらの洗練性の指標は英検2級から3級のモデル作文では有意な差が見られなかったことから、英検2級までは文法や語数が重視されていると考えられる。準1級と1級ではいずれも受験級が高いほど洗練された語彙が有意に多く使用されていた。これらの結果はいずれも、技能独立型の英作文における発表語彙は母語話者の心理言語学的な側面が重視されており、1級と準1級および準1級と

$< .001$ で高い受験級で有意に反応時間の長い語彙を使用しており、2級と準2級($p = .558$)、準2級と3級($p = .120$)では有意な差が見られなかった。この傾向は習得年齢についても同様であり、1級と準1級($p = .024$)、準1級と2級($p < .001$)でそれぞれ高い受験級で習得年齢の高い語彙が使用されていたが、2級と準2級($p = .101$)、3級と準2級($p = .648$)では非有意であった。この結果と図6を組み合わせて考えると、母語話者が5歳前後で受容語彙として習得する語彙は英検1級の発表語彙として求められており、大人の母語話者であっても6秒以上判断に迷う語彙は英検1級の発表語彙の特徴であると言える。しかし、これらの洗練性の指標は英検2級から3級のモデル作文では有意な差が見られなかったことから、英検2級までは文法や語数が重視されていると考えられる。準1級と1級ではいずれも受験級が高いほど洗練された語彙が有意に多く使用されていた。これらの結果はいずれも、技能独立型の英作文における発表語彙は母語話者の心理言語学的な側面が重視されており、1級と準1級および準1級と

2級で洗練性が異なることを示している。つまり、発表語彙の洗練性は熟達度を反映する指標であると主張する先行研究(e.g., Guo et al., 2013; Kyle & Crossley, 2016)と一致する結果が得られた。

語彙の洗練性が有意に高くなるのは英検の準1級以上であった。このことから、英検3級や2級を受験する中高生を指導している教育実践者は語彙の洗練性(言語的複雑性)よりも語数(流暢性)と文法の正確性を重視することにより、限られた授業時間の中で効率的にライティングを指導することができると考えられる。教育現場でのライティング指導では、表現する内容がなければ十分な量の作文をすることができないため、Weigle(2002)の分類における方略的能力がより重視されていると言える。例えば、作文のアイディアを生成するためのプレライティング活動として、ペア・グループでのブレインストーミング(スピーキング)や英語で書かれたウェブを使った情報の収集(リーディング)が行われる。これらの活動はライティングだけでなく他の技能を組み合わせた技能統合型活動である。続く調査2では、このような技能統合型タスクに着目して、その特徴を語彙の洗練性の観点から明らかにする。

4

調査2

4.1 目的

調査2では TAALES によって算出される洗練性指標に基づいて、技能独立型と統合型のタスクの違いと発表語彙の関係を明らかにする。RQs は以下の2点である。

RQ2-1

ライティングタスクの違いはどの洗練性指標によって予測されるか。

RQ2-2

ライティングタスクの違いによって発表語彙の特徴はどのように異なるのか。

4.2 方法

4.2.1 参加者

調査2に参加した協力者は、東京都内の私立大学、または茨城県の国立大学に通う大学生33名であった。協力者は3ヶ月程度の海外渡航経験はあるものの英語圏の教育機関に長期間留学してはいなかった。またアンケートによるプロフィールの調査の結果、英語熟達度は CEFR で B2 から C1 レベルであると推定された。

4.2.2 マテリアル

4.2.2.1 技能統合型タスク

統合型では技能間の結びつき(根岸, 2011)が重要な概念となる。調査2で対象とする技能統合型ライティングタスクは特定のトピックと英文が与えられ、書き手はそれをもとに自分の意見を述べるので、技能間の結びつきは弱いものを対象とする。この理由は、受容技能における熟達度の影響をできるだけ少なくするためにである。使用したソーステキストは「自動車の運転免許を取得する年齢を引き上げるべきか」というトピックで賛成意見と反対意見の両方を含む2つである。表3にテキストの特徴を示す。

4.2.2.2 技能独立型タスク

独立型タスクでは、統合型タスクと同様に「自動車の運転免許を取得する年齢を引き上げること」を題材に、賛成か反対か意見を述べるタスクを実施した。

■表3: 調査2で使用したソーステキストの特徴

難易度	語数	推定 CEFR-J	ARI	VperSent	AvrDiff	BperA	FKGL	S
賛成	112	B1.2	7.68 (A2.2)	3.43 (C1)	1.57 (B1.2)	0.13 (A2.2)	8.8	1844.19
反対	96	A2.2	5.81 (A1.3)	2.50 (B1.2)	1.58 (B1.2)	0.12 (A2.1)	6.1	3318.00

(注)読みやすさの指標等はCVLA ver1.1(Uchida & Negishi, 2018)にて算出。

4.2.3 手順

調査協力者は個別に2つのタスクをそれぞれ約20分で取り組んだ。いずれのタスクもZoomを通してコンピュータ上で行われた。作文中の画面は収録されており、書き手がリアルタイムに行っている推敲作業なども記録された。統合型タスクではソーステキストを協力者が印刷し、手元で参照しながらタスクが行われたが、辞書の使用は禁止された。協力者は最初に独立型に取り組み、その後統合型に取り組んだ。

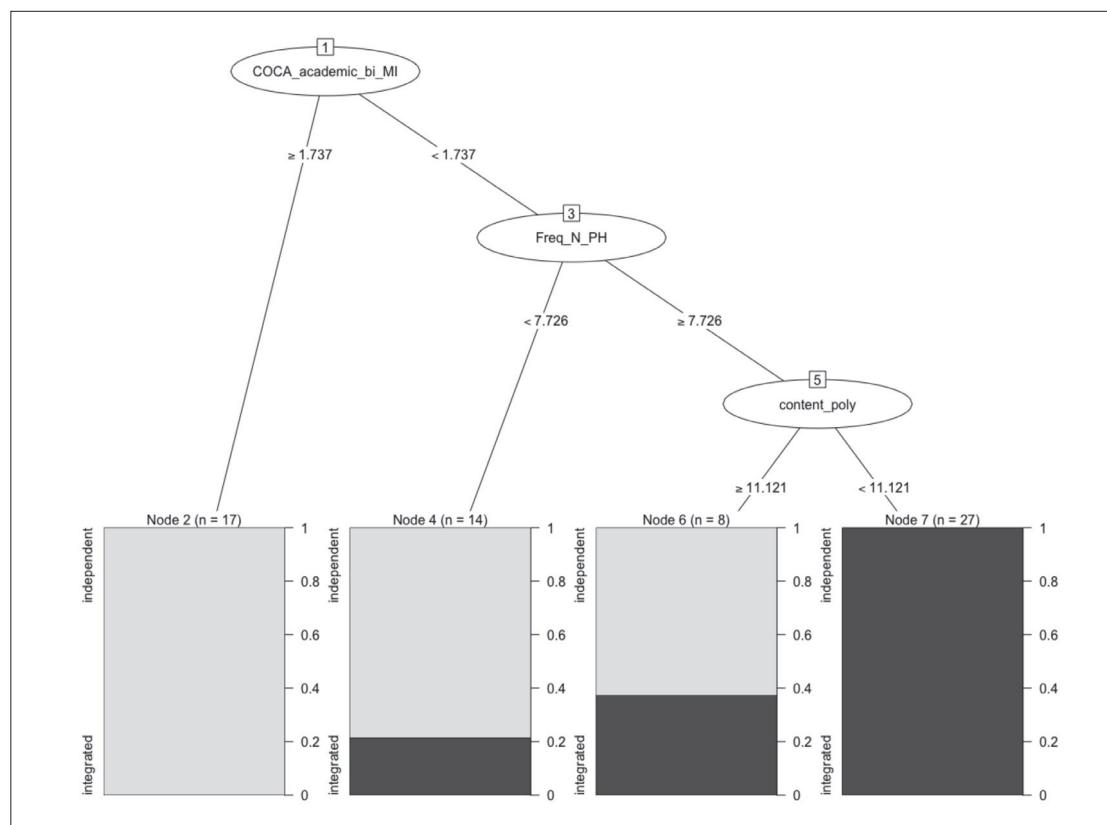
4.2.4. 採点と分析

調査1と同様にTAALESによるテキストの分析に先立ち、固有名詞や記号などは除去された。RQ1の焦点であるタスク分類を予測する洗練性指標を選出するために、説明変数をTAALESによって算出される洗練性の指標として決定木分析が行われた。その後、RQ2に解答するため、AntConcを用いてbigramおよびtrigramを抽出した。

4.3 結果と考察

4.3.1 RQ1:2種類の英作文に含まれる洗練性指標と分類の予測

TAALESによる洗練性指標算出の結果を表4、決定木分析の結果を図7にそれぞれ示す。2つのライティングタスクの分類を最も予測したのはCOCA Academic Bigram MI(アメリカ英語を収録したコーパスであるCOCAの学術サブコーパスにおける、2語の連続のn-gramの結びつきの強度)であった。技能独立型ではMIスコアの高いbigramを産出していると言える。第2のノードは音韻に関する近傍語頻度(近傍語頻度が大きいほど、形式的に類似した語彙が多く洗練性が高い)であった。そして第3のノードには内容語の多義性が含められた。技能統合型の作文の多くは内容語の語義が11よりも小さい語彙を使用し、技能独立型の作文では多義性の高い語彙を使用していることが示された。



■図7: 調査2の決定木分析の結果

■表4: 調査2の記述統計量

タスク	技能統合型		技能独立型	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
洗練性指標				
語数	134.79	32.49	122.94	43.74
COCA Academic Bigram Association Strength (MI)	1.44	0.17	1.66	0.23
COCA Academic Trigram Bigram to Unigram Association Strength (MI^2)	8.08	0.30	8.35	0.33
Phonological Neighborhood Frequency (homophones excluded)	7.89	0.14	7.84	0.18
Polysemy CW	10.14	1.06	9.95	1.95

第1ノードを適切に解釈することは難しいが、技能統合型タスクの方が独立型よりもMIスコアの高い連語を使用していることが示された。この結果は TAALES を用いて発表語彙を分析した Kyle and Crossley (2015, 2016) とは異なる。この理由として考えられるのは、認知資源の配分である。独立型タスクでは、未発達の連語よりもむしろ発表語彙として定着し自動化している単語や自由連結を使うことで、新たにアイディアを生成するための認知資源を節約していた。連語の発達も単語と同様に受容・発表の連続体であると捉えると (Meara & Bell, 2001), 未発達の項目を使用するには一定以上の認知資源を消費しなければならないと考えられる。統合型では新たにアイディアを生成することなくソーステキストのアイディアを利用することができるため、豊富な認知資源を形式面に配分することができた。それゆえ、発表語彙として未発達な連語を用いることができたと考えられる。

第2ノードは、音韻に関する近傍語頻度が7.726よりも大きければ統合型、小さければ独立型であることが多いことを示している。つまり、統合型タスクでは洗練性の高い語彙が用いられていたと言える。しかし近傍語頻度と反応時間の関係について検証した先行研究では、母語話者を対象とすることが多く、語彙認知の受容に関わる指標であることに注意が必要である。反応時間の研究では先行刺激が促進もしくは抑制の影響を与えることが示されているが (e.g., Andrews, 1989), 本研究でも統合型タスクのソーステキストで提示された語彙が発表語彙のメンタルレキシコンを活性化させた可能性がある。つまり形

式的に類似した語が刺激として提示されていれば、書記素や音韻に関するメンタルレキシコンを刺激し、ライティングでも使うことができるようになったと考えられる。ただし、作文中にソーステキストを参照することができたため、方略的な語彙使用があった可能性は十分に高く、読解時に提示されていた語彙が作文時にどのように使用されているかは検証の余地がある。

最後のノードである内容語の多義性は、英語語彙に関する大規模データベースである WordNet (Fellbaum, 1998) に基づいている。独立型では 11,121 個以上の語義を有している語が多いが、統合型では語義がそれよりも少なく洗練性の高い語彙が多いことを意味している。この第3ノードは、語彙の洗練性本来の意味で用いられる指標である。Laufer and Nation (1995) では洗練性の概念を使用頻度と対応づけて語彙頻度プロフィールを作成しているが「洗練された語彙」が意味するのは、その難しさである。洗練性の概念に基づくと、作文で複雑な表現を使うことができれば優れているとみなすため、平易な言葉で内容を伝えることに主眼を置くコミュニケーション活動では批判されることが多い。古くはルネサンス時代において洗練された語彙は「インク壺言葉」として批判されていた。ルネサンスでは、より専門的で多義性が低いようなラテン語やギリシア語由来の語彙を知識人が用い、多義性の高いゲルマン語由来の語彙は庶民が用いるものであった。インク壺言葉は知識人の難解な言い回しをインクの匂いがすると批判したものであるが、発表語彙の洗練性が高いほど言語熟達度が高いことを示していると言える。

母語話者であっても意見の分かれる語彙の洗

練性であるが、全く不要であるわけではない。例えば、計量的に作文を分析する上では、作文全体の長さに影響を受けてしまったり同じ長さを比較する際にも問題が生じたりする多義性を補完する役割を担っている。さらに、洗練性には頻度の観点だけでなく、調査2で示されたようなMIスコアや近傍語サイズ、内容語の多義性などの様々な指標が存在するため、多角的に発表語彙を捉えることができると言える。

4.3.2 RQ2:n-gramに基づく発表語彙の特徴

RQ2ではタスクの違いによる発表語彙の特徴を扱っているが決定木分析の結果から、単語だけではなく連語について検討することの重要性が示された。チャンクや定型表現と呼ばれる連語はしばしば熟達度の指標の1つとして使用される(e.g., Siyanova-Chanturia, & Pellicer-Sanchez, 2018)。母語話者や熟達度の高い学習者であれば、*by and large*という表現は個々の単語に分割せずに全体的に捉えて意味を解釈することができるが、初学者は単語を単語に分解して意味を捉えようとし、不透明な表現は正しく理解することができないことが多い。それゆえ、連語を発表語彙として習得していることは熟達度の指標の1つであるとみなすことができる。ここでは bigramに関する分析を AntConcによって bigram および trigram を抽出し、それぞれのタスクにおける発表語彙の特性を明らかにする。特徴語分析の結果、上位5項目は独立型と統合型でほとんど同じ n-gram が生成された(Appendix 1)。

TAALESで出力されたMIスコアが1.737以上のbigramとそれ以下の2つに分けると、独立型では *should, I, think, Japan, disagree, be* が特徴語として提示された。コンコーダンスに基づくと、*should* は主語に *government*を取り、(not) *raise* と共に起ることが多かった。このコロケーションはトピックに由来するものである。2つめの特徴語である *I* は *think, agree, disagree* と共に起し、*think* は *for this reason* や *so, therefore* に後続する一方、賛否について述べる場合には切り出し表現として機能していた。コンコーダンスプロットを見ると、*I think* は中盤から後半にかけて出現しており、まとめの表現として機能していると考えられる。一方、統合型の特徴語と

しては *should, disagree, Japan, we, who, too, can, think* が確認された。コンコーダンスラインにて、*should* は独立型と同様に、*government* を主語に選択し *raise* と共に起していたが、*we* を主語とする例も多く見られた。当該の *we* は *if* や *so* に導かれ、*must, should, can* などの助動詞と共に起していた。つまり、読み手を抱き込むようにした書き手の主張が強調されている。そして、*too* のクラスターとしては *young* や *danger*、そして不定詞の *to drive* が見られた。この *too A to do* の構文は独立型でも数件見られたが、ソーステキストで多く使用されていた。

ここでは *too A to do* の構文のようにスロットを埋めるプレハブ表現に着目する。スロットを提供するタイプの連語は全てを一から組み立てる必要がなく認知資源を大きく消費しないため、アイディア生成の必要がある独立型タスクで多く見られると予想される。しかし実際には、このタイプの連語は統合型の特徴語として確認された。この理由として、ソーステキストの存在が影響していると考えられる。具体的には RQ1 で得られた近傍語サイズの結果と同様に、読解によってメンタルレキシコンが活性化されたため、産出に至ったと考えられる。RQ1 では単語レベルの可能性であったが、連語においても活性化の可能性が示された。

両者に見られた *I think (that)* は MIスコアが3を超えており、書き手の意見を表す機能がある。統合型では作文の冒頭と中盤から後半で使用されることが多く、さらに賛成・反対の立場を述べた上で *and* や *but* といった接続詞に後続して *I think* が用いられていた。この理由はソーステキストが存在する場合に、自分の意見とテキストで示されている事実や意見を区別するためである。本調査の協力者はアカデミックライティングなどの授業を受講している大学生であり、剽窃に関する十分な知識を有しており、立場を表す機能表現を発表語彙知識として有していた可能性がある。

4.4 調査2のまとめ

調査2では2種類のライティングタスク(技能独立型と技能統合型)を比較し、TAALESによって

算出される語彙の洗練性という観点から発表語彙の特徴を検証した。決定木分析の結果、タスクの分類を予測する洗練性の指標として、bigramにおけるMIスコア、音韻の近傍語サイズ、内容語の多義性の3つが挙げられた。また発表語彙の中でもbigramに焦点を当てた結果、その使用にはソーステキストの存在が影響している可能性が示された。

5

結論

本研究は2つの調査から構成されていた。調査1は英検のモデル英作文を分析し、2級以下では洗練性に有意な差は見られなかったことから、教育現場では英作文の正確性と流暢性に焦点を当てる必要性を示した。調査2では2種類のタスクが発表語彙に与える影響について検証した。調査1と異なる指標が分類を予測していたが、この理由はソーステキストの存在に起因すると考えられる。技能統合型タスクは真正性が高く、特に中学校での指導では統合型タスクを行うように教科書が設計されており、近年の4技能統合型の指導で広く教育実践現場で取り入れられる。しかし、タスクの遂行には一定以上の熟達度が求められる点(Plakans, 2013)、技能統合の強度(Cumming, 2014; 根岸, 2011)が高い場合には受容技能を活用する段階で問題を抱えてしまう学習者がいる可能性に留意しなければならない。それゆえ、受容の段階ではジグソーやディクトコンポといった協働学習を取り入れ、足場かけを提供しながら受容に関するタスクを乗り越え、発表タスクに問い合わせるという方法が有効であるだろう。

2つのライティングタスクの分類に最も寄与していた洗練性指標はn-gramの結びつきの強さであった。この結果には調査協力者がこれまでに受けてきた語彙指導が関係している。本調査の協力者は日本の中学校・高校で6年間英語を学習した大学生であった。連語表現の指導は主にディスコースマーカーを用いたリーディングで扱われるが、現在の中高でのライティング教育ではword boxにてスロットを埋める形式の定型表現を足場かけとして提示することが多い。言語使用を通じた付隨的な学習法は大量に存在する連

語を学習する上で形式・意味・機能を結びつけるのに効果的であり(Nation, 2013)、受容技能を通じた付隨的語彙学習よりも発表技能を通じた付隨的学習の方が高い関与負荷を有するため効果的である(Laufer & Hulstijn, 2001)。しかし付隨的語彙学習はそもそも効率がよいわけではない(Webb, 2008)。現在の中高での連語の指導が受容において明示的に、発表において暗示的に指導されることが多いため、連語表現が発表語彙として未発達であったと考えられる。そこで連語を指導する際には、高頻度語に焦点を当てて学習することができ、学習の効率の高い意図的な方法と組み合わせることが肝要となる(Nation, 2013)。例えば、大学でのアカデミックライティングにおいて、分野に高頻度な専門用語を含めた連語表現を中心とした明示的な方法で指導すると、学習効率を高めることができると考えられる。

本研究では言語の複雑性を表す語彙の洗練性に着目したが、この指標は英検1級から準1級レベルの高い熟達度を有する大学生や大学院生の学術的な文章を評価する際に特に有効である。学術目的の英語では分野における低頻度の専門用語が繰り返し使用される傾向があるため、頻度の観点だけでは発表語彙を妥当に評価することは難しい。そこで英語のレポートの採点では、TALESによる洗練性指標を評価観点の1つに用いることができる。その際、心理言語学的な反応時間や習得年齢を参考にすることが効果的であることが本研究から示唆された。

本研究の限界点は大きく3つある。1つは作文のトピックに関する点である。間接的な発表語彙の測定では題材に影響を受けてしまうため複数のトピックを用いるべきであるが、本研究では1つのみであった。今後は複数のトピックを用いた調査が必要である。2つめは調査サンプルに関する点である。本調査の参加者は大学生のみを対象としており英語熟達度が限定的であるため、日本の英語学習者に結果を一般化することは難しい。最後にTALESの問題点に関して、膨大な数の指標が算出されるため詳細な分析を行える点では有益であるが、使用されているコーパスの特徴や指標の意味が分かりにくいものも多い。ユーザーフレンドリーなツールとしては頻度に基づいたSが挙げられるが、洗練性の様々な側面

を考慮した直感的に分かりやすいツールの開発 | が期待される。

Appendix 1 AntConcによって抽出された調査2のn-gram

ランク	bigram項目(頻度)		trigram項目(頻度)	
	統合型	独立型	統合型	独立型
1	driving age (74)	the driving (64)	the driving age (64)	the driving age (58)
2	the driving (64)	driving age (62)	raise the driving (34)	raise the driving (32)
3	in Japan (50)	in Japan (58)	age in Japan (26)	age in Japan (26)
4	raise the (36)	car accidents (44)	driving age in (22)	driving age in (24)
5	I think (36)	young people (44)	should raise the (16)	cause car accidents (18)

謝辞

本研究を発表する貴重な機会を与えてくださいました公益財団法人 日本英語検定協会と関係者の皆様に厚く御礼申し上げます。特に助言者である寺内一先生には、非常に有益なご指導をいた

だき、大変感謝しております。そして、筑波大学の卯城祐司先生、共に学ぶ神村幸蔵さん、小木曾智子さん、佐々木大和さん、前田啓貴さんには多くの示唆とサポートをいただきました。本研究にご協力いただいた皆様に深くお礼申し上げます。

参考文献

(*は引用文献)

- Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 802-814. <https://doi.org/10.1037/0278-7393.15.5.802>.
- Baba, K. (2009). Aspects of lexical proficiency in writing summaries in a foreign language. *Journal of second language writing*, 18(3), 191-208. <https://doi.org/10.1016/j.jslw.2009.05.003>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Beglar, D., & Nation, P. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13. https://www.lextutor.ca/tests/nation_beglar_size_2007.pdf
- Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, 37(5), 639-668. <https://doi.org/10.1093/applin/amu059>
- Fellbaum, C. (1998) *WordNet: An Electronic Lexical Database*. MIT Press.
- Conklin, K. (2019). Processing single-word and multiword items. In S. Webb, (Ed.). *The Routledge Handbook of Vocabulary Studies*, (pp. 174-188). Routledge.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238. <https://doi.org/10.2307/3587951>
- Crossley, S. A., & Skalicky, S. (2019). Examining lexical development in second language learners: An approximate replication of Salsbury, Crossley & McNamara (2011). *Language Teaching*, 52(3), 385-405. <https://doi.org/10.1017/S0261444817000362>
- Cumming, A. (2014). Assessing integrated skills. In A. Kunnan (Ed.), *The companion to language assessment* (vol. I : Abilities, contexts, and learners; Part 2: Assessing abilities), (pp. 216-229). John Wiley & Sons.
- Gebril, A. (2009). Score generalizability of academic writing tasks: Does one test method fit it all? *Journal of Language Testing*, 26(4), 507-531. <http://dx.doi.org/10.1177/0265532209340188>.
- Gebril, A., & Plakans, L. (2016). Source-based tasks in academic writing assessment: Lexical diversity, textual borrowing and proficiency. *Journal of English for academic purposes*, 24, 78-88. <https://doi.org/10.1016/j.jeap.2016.10.001>
- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, 29(2), 228-244. [https://doi.org/10.1016/0749-596X\(90\)90074-A](https://doi.org/10.1016/0749-596X(90)90074-A)
- Granger, S. (2019). Formulaic sequences in learner corpora: Collocations and lexical bundles. In Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (Eds.), *Understanding formulaic language a second language acquisition perspective*, (pp. 228-247). Routledge
- Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*, 13(4), 403-437. <https://doi.org/10.1075/ijcl.13.4.02gri>
- Guo, L., Crossley, S. A., & McNamara, D. S. (2013). Predicting human judgments of essay quality in both

参考文献 (*は引用文献)

- integrated and independent second language writing samples: A comparison study. *Assessing Writing*, 18(3), 218–238. <https://doi.org/10.1016/j.asw.2013.05.002>
- Hunston, S. (2002). *Corpora in applied linguistics*. Ernst Klett Sprachen.
- Ishikawa, S., Uemura, T., Kaneda, M., Shimizu, S., Sugimori, N., Tono, Y., ... Murata, M. (2003). *Japan Association of College English Teachers (JACET) 8000: JACET list of 8000 basic words*. JACET.
- Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning*, 63, 87–106. <https://doi.org/10.1111/j.1467-9922.2012.00739.x>
- Kim, M., Crossley, S. A., & Kyle, K. (2018). Lexical sophistication as a multidimensional phenomenon: Relations to second language lexical proficiency, development, and writing quality. *The Modern Language Journal*, 102(1), 120–141. <https://doi.org/10.1111/modl.12447>
- Kojima, M., & Yamashita, J. (2014). Reliability of lexical richness measures based on word lists in short second language productions. *System*, 42, 23–33. <https://doi.org/10.1016/j.system.2013.10.019>
- Kyle, K. (2019). Measuring Lexical Richness. In S. Webb, (Ed.), *The Routledge Handbook of Vocabulary Studies*, 454–476. Routledge.
- Kyle, K., & Crossley, S. (2016). The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, 34, 12–24. <https://doi.org/10.1016/j.jslw.2016.10.003>
- Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *Tesol Quarterly*, 49(4), 757–786. <https://doi.org/10.1002/tesq.194>
- Laufur, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322. <https://doi.org/10.1093/applin/16.3.307>
- Laufur, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33–51. <https://doi.org/10.1177/026553229901600103>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474–496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, 27, 57–86. doi:10.1177/0741088309351547
- Meara, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16(3), 5–19. <http://www.lognistics.co.uk/vlibrary/meara&bell2001.pdf>
- Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I.S.P. (1983). Testing and teaching vocabulary. Guidelines, 5, 12–25. <https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/paul-nations-publications/publications/documents/1983-Testing-and-teaching.pdf>
- 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. (2013). Assessment of integrated skills. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. (pp. 205–212). John Wiley & Sons.
- Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing*, 17(1), 18–34. <https://doi.org/10.1016/j.asw.2011.09.002>
- Polio, C., & Friedman, D. A. (2016). *Understanding, evaluating, and conducting second language writing research*. Taylor & Francis.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487–512. <https://doi.org/10.1093/applin/amp058>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2018). Introduction. Formulaic language: Setting the scene. In Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (Eds.), *Understanding formulaic language a second language acquisition perspective*. (pp. 1–15). Routledge.
- Uchida, S. and M. Negishi (2018) Assigning CEFR-J levels to English texts based on textual features. In Y. Tono and H. Isahara (eds.) *Proceedings of the 4th Asia Pacific Corpus Linguistics Conference (APCLC 2018)*, (pp. 463–467). <http://hdl.handle.net/2324/2244112>
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20, 232–245. <http://nflrc.hawaii.edu/rfl/October2008/webb/webb.pdf>
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics*, 168(1), 33–69. <https://doi.org/10.1075/itl.168.1.02web>
- Weigle, S. C. (2002). *Assessing writing*. Ernst Klett Sprachen.
- 相澤一美・望月正道. (2010). 『英語語彙指導の実践アイディア集・活動例からテスト作成まで』. 大修館書店
- 根岸雅史 (2011). 「技能統合の評価をどうするか」. 『英語教育』第60巻2号 (5月号), 29–31. 大修館書店。
- 佐藤敬典. (2018). 「テスト作成に必要な理論を学ぼう」. 小泉利恵, 印南洋, 深澤真. (編著). 『実例でわかる 英語テスト作成ガイド』. (pp. 76–79). 大修館書店.
- 望月昭彦. (2015). 「ライティングと他技能の統合型的活動の評価」. 望月昭彦, 深澤真, 印南洋, 小泉利恵. (編著). 『英語4技能評価の理論と実践 —Can-Do・観点別評価から技能統合的活動の評価まで』. (pp. 92–101). 大修館書店.

第32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

自動英文解析ツールを用いた英作文採点の 妥当性検証:Coh-Metrix と Text Inspector の 指標に基づいて

研究者:茨城県／筑波大学大学院 在籍 佐々木 大和

《研究助言者:村木 英治》

概要

本研究では、英文解析ツールである Coh-Metrix と Text Inspector の自動英作文評価への利用可能性について2つの調査を行った。まず、調査1では、Coh-Metrix と Text Inspector を用いて、英検が公表している1級から3級の英作文の模範解答例のテキスト分析を行い、テキスト特性を検証した。結果として、受験級が高くなるほど、トピックが難しくなり、幅広い単語や時制・相を用いた難易度が高い英作文を書くように求められていることがわかった。次に、調査2では、Coh-Metrix と Text Inspector を用いて、実際の日本人英語学習者の英作文のテキスト分析を行い、算出された指標の傾向と英語力との関係、英作文得点との関係を調査した。結果として、英作文のテキスト特性に関しては、受験級が上がるにつれて、英作文の馴染みやすさや時制や相の利用、難易度、語の多様性が増加する傾向にあることがわかった。一方、求められる語数が多くなるほど、簡単な馴染みのある文や同じ動詞を繰り返し使うことがわかった。算出された指標と学習者の英語力の関係に関しては、求められるトピックや語数が異なることで、英語力を予測する指標が変わる可能性が示唆された。算出された指標と英作文評価との関係性については、受験級ごとに英作文評価と関係のある指標が異なっていたが、全ての受験級を通して、語の長さや頻度、親密度、多様性のような指標が予測変数に含まれていたことから、語に関する指標が自動英作文評価に影響を与えている可能性があることが示唆された。

1 はじめに

近年、大規模英語テストを入試に利用する学校が増えている。リーディングやリスニングテストに関しては、多肢選択式問題が取り入れられ、採点の妥当性が確保されている。一方、ライティングにおいては、採点前トレーニングを行ったり、熟練した採点者を確保したりしても、多肢選択式問題に比べ、主観性が介在してしまうため、採点者間でブレが生じる可能性がある。また、普段の英語の授業においても、学習者の英語ライティング能力を測定するために、教師が英作文を課す場合が多い。しかし、教師にとって、この英作文を採点・評価する作業というのは、かなりの労力と時間を要する。そこで、近年、英作文の自動評価が注目を浴びてきている。一方、英作文自動評価ツールには有料のものも多く、手を出しにくい場合も多い。

以上の背景を踏まえ、本研究では、近年、読解研究において、ウェブベースの無料の自動英文解析ツールとして活用されている Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004) と Text Inspector (Bax, 2012) の英作文自動評価への応用可能性を検証する。このツールは自然言語処理とデータマイニングを利用し、さまざまな指標を用いて、英文の言語的特徴を客観的に評価することが可能である。

2

先行研究

2.1 ライティング

ライティングは、書き手が持つ文章を書くための技能や知識を利用する認知プロセスでもあり、ある特定の目的や読み手を伴う場面において行われる状況内の活動(situated activity)である(Polio & Friedman, 2017)。ライティングは文字を使って、相手に情報などを伝える活動であり、インターネット等で世界中の人々と繋がる可能性が高まる現代社会において、英語で情報を伝えるニーズは高まっている(佐野, 2013)。しかし、平成30年に告示された高等学校学習指導要領(文部科学省, 2018)において、「書くこと」の言語活動が適切に行われていないことが指摘され、その問題を踏まえ、話すことや書くことによる発信能力の育成を目的とした「論理・表現」という外国語科目が新設された。

「1.はじめに」で述べたように、英語のテストにおいて、英語ライティングの能力を測定するために、生徒に英作文を書かせる場合が多い。また、授業においても、英語ライティング向上のために、英作文課題を課す教員が多い。しかし、その英作文の採点や評価には、かなりの労力と時間がかかり、教員の負担となっている。また、英作文の評価には、評価者の主觀が介在する可能性が高く、評価者トレーニングを行うなど妥当性の確保が難しい。そこで、ツールを使った英作文の自動評価を利用する教員やテスト会社が増えている。

2.2 英作文の評価法と自動評価

英作文の評価法としては、総合的評価(holistic scoring)と分析的評価(analytic scoring)が挙げられる。前者は、評価者が英作文全体を読んで、点数を1つ与える評価法であり、後者は、初めにいくつか観点を定め、その観点ごとに点数を与える評価法である。英作文の自動評価には、総合的評価が用いられることが多い(小林・金丸, 2012; Shermis & Burstein, 2003)。

英作文の自動採点(automated essay evaluation)とは、コンピュータ・プログラムを通して、作文を評価したり、採点したりすることである。近年、Educating Testing Serviceのe-rater Scoring Engineや日本英語検定協会のWriting Tutorなど、様々な機関が英作文の自動採点にコンピュータ・プログラムを利用している。しかし、これらの自動評価は教員や学習者がいつでも使えるものではなく、有料の場合も多い。そこで、本研究では、無料のウェブベースの英文解析ツールであるCoh-MetrixとText Inspectorの英作文自動評価への利用可能性を検証する。

2.3 英文解析ツール:Coh-MetrixとText Inspector

Coh-MetrixやText Inspectorといった英文解析ツールは、さまざまな指標を用いて、英文の言語的特徴を客観的に評価することができるものである。Coh-Metrixは英文の結束性(文章中の要素の結びつき)を評価できる点が特徴的であり、複数の関連する指標を用いて、Referential Cohesion, Verb Cohesion, Deep Cohesionといった英文の結束性の指標を算出することが可能である。特に読解研究において、幅広く利用されてきている(McNamara, Graesser, McCarthy, & Cai, 2014)。

それに対し、Text Inspectorは英文に対して、ヨーロッパ言語共通参照枠(Common European Framework of Reference for Languages; CEFR)に基づいたScorecardを算出することができる。このScorecardでは、Percentage(100%が英語母語話者レベルを表す), Number of Metrics Used(参考にした指標の数), CEFR Levelを確認することができる。また、このScorecardは、リーディング、ライティング、リスニングといったテキストの種類によって、参考にする指標を変えて算出されている。

以上のような英文解析ツールをライティングに応用した研究がなされている。Coh-Metrixを用いた研究に関しては、特に、Coh-Metrixで算出された指標と評価者による英作文評価の関係性を検証している研究が多い(e.g., Crossley & McNamara, 2010; 2012; McNamara, Crossley, & McCarthy, 2010; Guo, Crossley, &

McNamara, 2013; Zedulius, Millis, & Schooler, 2019)。母語話者の英作文を対象とした研究では、英作文の語彙の洗練さ (lexical sophistication) や統語的複雑さ (syntactic complexity) の指標が採点者による評価に影響を与え、Coh-Metrix の特徴である結束性に関する指標はあまり影響を与えていない (McNamara et al., 2010) という結果や、評価者による結束性・一貫性に関する評価にも、結束性に関する指標は関連していない (Crossley & McNamara, 2010) という結果が得られている。また、英作文の創造性に着目すると、英作文の馴染みやすさ (narrativity) や語や内容の重複 (referential cohesion), 語の具象性 (word concreteness) などの言語的特徴が採点者による評価に影響を与えていることが示唆されている (Zedulius et al., 2019)。一方、外国語としての英語を学ぶ学習者の英作文を対象とした研究では、英作文の語の多様性 (lexical diversity), 語の頻度 (word frequency), 語の意味内容 (word meaningfulness), 相の繰り返し (aspect reputation), 語の親密度 (word familiarity) が採点者による評価を予測するという結果 (Crossley & McNamara, 2012) や、タスクの種類が異なっていても、英作文の長さなどの言語的特徴が採点者による評価を予測するという結果 (Guo et al., 2013), 語彙に関する指標が採点者による評価と関係するという結果 (Aryadoust & Liu, 2015) が得られている。

Coh-Metrix で算出された指標と書き手の熟達度 (i.e., 英作文の質) の関係に関しては数はあまり多くはないが、研究が行われている (小林・金丸, 2012; Latifi & Gierl, 2020)。母語話者を対象とした研究では、複数のトピックの英作文間で比較しており、語や内容の重複 (referential cohesion) と語の多様性 (lexical diversity) に関する特性が英作文の質の評価に最も影響を与えているが、英作文のレベルやトピック、書き手の熟達度、利用したルーブリックなどにより影響を与える指標が異なる (Latifi & Gierl, 2020) という結果が得られている。また、外国語として英語を学ぶ学習者を対象とした研究に関して、日本人英語学習者を対象に行った小林・金丸 (2012) では、e-rater によって評価された書き手の熟達度を推定するにあたって、異なり語数の値が最も熟達度

の推定に寄与していることが示された。

Text Inspector を用いた研究については、数が限られている。Bax, Nakatsuhara, and Waller (2019) では、文章の繋ぎりを作る談話標識といった metadiscourse marker に着目し、学習者を対象に研究を行った。結果として、熟達度の高い書き手は熟達度の低い書き手よりも metadiscourse marker を使わないが、幅広く使用するということが示された。

2.4 本研究の概要と目的

以上のような先行研究を踏まえ、本研究では、外国語として英語を学ぶ日本人を対象にし、Coh-Metrix や Text Inspector といった自動英文解析ツールを用いた英作文自動評価の妥当性を検証する。調査1では、英検が公表している英作文問題の模範解答例のテキスト特性を Coh-Metrix と Text Inspector を用いて検証する。次に調査2では、日本人英語学習者の英作文を対象に、Coh-Metrix と Text Inspector で算出されたテキスト特性の傾向、英語力、英作文得点との関係を検証する。

3 調査1

3.1 目的

調査1では、英検が公表している英作文問題の模範解答例のテキスト特性を Coh-Metrix と Text Inspector で分析することで、各級において求められている英作文の特性を明らかにすることを目的とする。具体的には、1級から3級までの英作文問題の模範解答例を対象とし、異なる受験級において、どのようなテキスト特性が見られるかを明らかにする。検証課題 (Research Question: RQ) は以下のとおりである:

RQ1-1 Coh-Metrix や Text Inspector により算出される指標について、英検1級から3級の英作文模範解答例の特性はどうになっているか。

3.2 方法

3.2.1 マテリアルの収集

英作文問題の模範解答例のテキスト特性を明らかにするため、英検1級から3級までの模範解答例を収集した。具体的には、英検3級に英作文問題が導入された2017年度第1回から2019年度第3回までの英作文問題の模範解答例を各受験級(1級、準1級、2級、準2級、3級)からそれぞれ9テキスト、計45テキスト収集した。

3.2.2 分析指標

本調査で使用した指標はCoh-Metrixから9つ(Readability, Narrativity, Syntactic Simplicity, Word Concertedness, Referential Cohesion, Deep Cohesion, Verb Cohesion, Connectivity, Temporality), Text Inspectorから4つ(Percentage, Gunning Fog, Metadiscourse [Type], Lexical Diversity)である。

Coh-Metrixにおける指標に詳しい説明は以下のとおりである:

Readability:

テキスト中の内容語の重複や統語的類似性、語の頻度を考慮した読みやすさの指標。

Narrativity:

テキストがどれだけ日常的であるか、つまり馴染み深いかを示す指標。

Syntactic Simplicity:

テキスト中の文がどれだけ少ない語で、かつ簡単な馴染みのある統語構造が使用されているかを示す指標。

Word Concreteness:

テキスト中の内容語に具体的な語がどの程度含まれているかを示す指標。

Referential Cohesion:

文間やテキスト全体で重複する語や考えがどの程度含まれているかを示す指標。

Deep Cohesion:

テキスト中に因果的もしくは論理的な接続語がどの程度含まれているかを示す指標。

Verb Cohesion:

テキスト中に動詞の重複がどの程度含まれているかを示す指標。

Connectivity:

テキスト中に逆説・付加・比較の接続語がどの程度含まれているかを示す指標。

Temporality:

テキスト中に時間的な手がかりがどの程度含まれているか、時制や相がどの程度一致しているかを示す指標。

また、Text Inspectorにおける指標の説明は以下のとおりである:

Percentage:

100%が、母語話者が書く高いレベルのアカデミックな英作文を示す。

Gunning Fog:

1文あたりの平均単語数と長い綴りの単語を考慮した英作文の読みやすさの指標。

Metadiscourse (type):

接続詞といった談話標識がどの程度使用されているかを示す指標。

Lexical Diversity:

語の多様性、つまりどの程度幅広い語が使用されているかを示す指標。

3.3 手順

テキスト分析ツールのCoh-Metrix (<http://www.cohmetrix.com/>)とText Inspector (<http://www.textinspector.com>)を用いて、英検が公表している英作文問題の模範解答例を分析した。

3.4 結果と考察

表1はCoh-Metrixで算出した指標を受験級ごとに集計した記述統計である。結果より、級が上がるにつれて、Narrativityの値が減少傾向にあることがわかる。級が上がるにつれて、テキストの馴染みやすさが減少することから、受験者にとって難しいトピックが扱われていることがわかり、特に準1級と2級間の差は大きい。また、Temporalityにおいても、同様の減少傾向が見られる。級が上がるにつれて、時間を示すような語句を使わなくなったり、英作文に含まれる時制(現在形や過去形)や相(完了形や進行形)が多様化したりする傾向にあることがわかる。さらに、

Readability も準2級から1級にかけて減少傾向が見られ、難易度の高い英文を書くように求められていることがわかる。3級の値が準2級とあまり変わらないことに関しては、語数が少ないため、内容語の重複が少ないと起因すると考えられる。

一方、統語的な簡単さを示す Syntactic Simplicity や語の具体性を示す Word Concreteness は、級が上がるにつれて減少すると予測していたが、一貫した結果は得られなかつた。Syntactic Simplicity に関しては、3級が最も値が低いが、これは語数が少ないために値が低く出てしまった可能性がある。準1級と2級は類似した値が算出され、1級で値が減少している

ため、文中の語数が増え、また難しい統語構造を使うことを求められていることがわかる。Word Concretenessにおいては、3級が最も高く、準2級、2級、準1級の値は似ており、1級で減少している。1級では、難しいトピックが扱われる傾向にあることから、英作文で使用する語の抽象度が上がると考えられる。

Coh-Metrix の特徴である結束性・一貫性の値に関しては、級ごとで一貫した結果は得られていないが、どの指標においても1級の値が最も小さい。求められる語数が上がるにつれて、相対的に結束性や一貫性を示す接続語が少なくなり、英作文の結束性・一貫性が低くなる可能性がある。

■表1: Coh-Metrixによる英作文模範解答の級ごとの特徴

	1級 (n=9)		準1級 (n=9)		2級 (n=9)		準2級 (n=9)		3級 (n=9)	
	M	SD	M	SD	M	SD	M	SD	M	SD
Readability	11.37	2.33	16.60	3.05	26.55	5.42	34.85	3.74	32.82	8.40
Narrativity	14.49	5.98	28.90	15.94	68.46	15.11	86.13	5.98	88.86	22.04
Syntactic Simplicity	57.42	18.33	65.12	18.86	61.43	15.99	88.87	9.89	53.50	28.72
Word Concreteness	33.45	19.28	50.79	20.11	52.99	27.60	50.26	29.25	82.47	17.64
Referential Cohesion	19.82	18.93	30.59	23.94	70.89	26.03	66.59	30.27	83.78	25.03
Deep Cohesion	73.57	16.32	92.61	8.53	81.79	20.77	98.15	2.91	80.33	28.81
Verb Cohesion	20.45	14.80	36.24	23.65	83.41	13.33	68.94	33.51	49.35	28.59
Connectivity	11.06	8.37	1.61	1.73	36.98	33.52	9.27	13.05	28.23	23.95
Temporality	17.23	18.80	22.12	26.03	40.01	33.96	44.38	26.69	64.33	37.22

次に、表2にText Inspectorで算出した指標を受験級ごとに集計した記述統計を示す。表より、級が上がるにつれて、Gunning Fog, Lexical Diversity の値が上昇していることがわかる。Gunning Fogに関しては、英作文を読む際に、読み手がどの程度の能力が必要かを示す指標である。Coh-Metrixの結果と同様、級が上がるにつれて、難しい英作文を書くように求められていることがわかる。また、Lexical Diversityに関しては、級が上がるにつれて、様々な種類の語を使うように求められていることがわかり、2級と準1級の間の値の差が大きい。

Percentage は母語話者らしいアカデミックな

英作文が書けているかを表す指標であり、Text Inspectorにおいて、英作文の得点を表す。準2級から1級にかけて、得点が上昇傾向にあり、母語話者らしい英作文を書くように求められていることがわかる。また、他の指標と同様、準1級と2級の間に大きな値の差がある。

一方、Metadiscourseに関しては、減少傾向にある。これは様々な談話標識の数を表すものであり、接続詞など結束性・一貫性に関わる指標である。Coh-Metrixでの指標において、結束性・一貫性に関しては一貫した結果は得られず、1級が最も値が低くなるという結果であったが、Text Inspectorでは、一貫して減少していることがわかった。

表2: Text Inspectorによる英作文模範解答の級ごとの特徴

	1級 (n=9)		準1級 (n=9)		2級 (n=9)		準2級 (n=9)		3級 (n=9)	
	M	SD	M	SD	M	SD	M	SD	M	SD
Percentage	77.69	4.58	70.69	3.35	48.20	6.60	37.09	7.85	46.72	9.42
Gunning Fog	16.76	1.10	15.06	1.54	9.45	0.92	7.60	1.74	7.43	2.59
Metadiscourse	9.03	2.08	10.71	2.78	14.65	3.03	17.05	4.52	20.79	3.93
Lexical Diversity	111.83	23.55	102.28	20.44	68.78	15.96	56.17	7.92	n/a	n/a

3.5 調査1のまとめ

調査1では、英作文問題の模範解答例のテキスト特性を明らかにするために、1級から3級までの英作文模範解答例を Coh-Metrix と Text Inspector を用いて分析を行った。結果より、受験級が高くなるほど、トピックが難しくなり、幅広い単語や時制・相を用いたテキスト難易度が高い英作文を書くように求められていることが示唆された。

調査2では、実際に日本人英語学習者が書いた英作文を対象にし、Coh-Metrix や Text Inspector を用いてテキスト分析を行い、学習者の英語力とどのような関係があるのかを調査した。

4 調査2

4.1 目的

調査2では、日本人英語学習者の英作文を Coh-Metrix と Text Inspector を用いて分析し、算出された指標の特徴やその指標と学習者の英語力にどのような関係があるのかを調査する。また、Text Inspector で算出される英作文の得点(Percentage)が他の Coh-Metrix や Text Inspector の指標とどのような関係があるのかを検証する。検証課題(RQs)は以下のとおりである：

RQ2-1

学習者の英作文における、Coh-Metrix と Text Inspector で算出された指標は級ごとにどのような傾向があるのか。

RQ2-2 学習者の英作文における Coh-Metrix と Text Inspector で算出された指標は学習者の英語力とどのような関係があるのか。また、級ごとで異なるのか。

RQ2-3 Text Inspector で算出された英作文の得点(Percentage)は他の Coh-Metrix や Text Inspector の指標とどのような関係があるのか。また級ごとで異なるのか。

4.2 方法

4.2.1 協力者

私立大学に通う日本人英語学習者30名が調査に参加した。協力者は全員大学1年生であり、専攻は人文系であった。協力者は全員日本語母語話者で、少なくとも6年以上日本の教育機関で英語を学んでいた。自己申告の英語資格のアンケートによると、協力者は英検3級から2級程度であり、CEFRのA1からB1レベルであると推察される。実験は3日間に分けて行われたため、全ての実験を完遂した22名を分析対象とした。

4.2.2 マテリアル

(1) 英作文課題

英検の過去問題より、3級、準2級、2級から1つずつトピックを選定した。選定したトピックは次の表3のとおりである。受験級に関しては、大学1年生を高校卒業程度とし、2級までとした。指定語数はそれぞれの級にならい、3級のトピックは25語から35語、準2級は50語から60語、2級は80語から100語とした。指示は実際の英検の出題

形式に則り、「以下の質問について、あなたの考え方とその理由を2つ英文で書きなさい」とし、2級のトピックではポイントを3つ(Convenience, Cost, The environment)与えた。解答は紙ベー

スで行われた。協力者の英作文の語数の平均は、3級では35.45語($SD = 10.13$)、準2級では39.73語($SD = 17.38$)、2級では61.32語($SD = 26.05$)であった。

■表3: 英作文の受験級とトピック

受験級	英作文のトピック
3級	What day of the week do you like the best?
準2級	Do you think it is important for children to play sports?
2級	It is often said that people today use too much electricity. Do you agree with this opinion?

(2) 英語力測定課題

協力者の英語力を測定するために、英検IBAを用いた。リーディングとリスニングの2技能を測定するものを使用し、2級から3級を測定することが可能なテストBを使用した。形式は英検と同様のもので、すべて多肢選択式問題であった。リーディング問題は語彙・文法問題20問、会話文問題5問、長文問題10問で構成されており、リスニング問題は会話文問題15問、英文問題15問で構成されている。

4.2.3 手順

調査は3日間に分けて一斉に実施された。全体の所要時間は95分程度であった。1日目と2日目に協力者は英作文課題を行った。1日目に英検3級と準2級のトピック、2日目に英検2級のトピックに関する英作文を書き、それぞれ25分間で行った。この英作文課題の際、協力者は実際の試験を想定し、辞書等で単語を調べることはできなかった。

3日目に英検IBAテストが実施された。リーディングが25分、リスニングが20分の計45分で行われた。この際、英作文課題と同様、辞書等の使用は禁止されていた。

4.2.4 採点・分析

英作文課題に関しては、調査1と同様、ウェブベースのテキスト分析ツールであるCoh-MetrixとText Inspectorを用いて、学習者が取り組んだ各級の英作文を分析した。Coh-Metrixでは、綴りの誤りを分析することができないので、綴り

の誤りに関しては、調査者が書き起こす際に修正した。また、英作文問題の模範解答例では、3級、準2級、2級と1段落で構成されているため、協力者の英作文も1段落に統制を行った。

英語力測定課題に関しては、日本英語検定協会で採点がされ、得点は成績表のCSEスコアを参考にした。満点スコアは1300点である。

RQ2-1に関して、研究1と同様の分析指標を用いて、記述統計を算出し、受験級ごとの傾向を確認した。

RQ2-2に関して、Coh-MetrixやText Inspectorで算出される指標が膨大であるため、Crossley and McNamara(2012)を参考に、英語力と算出された指標で各級ごとの英作文で相関分析を行い、相関係数が大きい指標を選定し、変数の絞り込みを行った。絞り込まれた変数を独立変数、英検IBAテストで測定された英語力を従属変数に重回帰分析を行った。

RQ2-3に関しても、RQ2-1と同様の手順で変数の絞り込みを行った。絞り込まれた変数を独立変数、Text Inspectorで算出された英作文の得点(Percentage)の値を従属変数に重回帰分析を行った。

4.3 結果と考察

(1) 英語力測定課題

表4は英検IBAテストの記述統計である。

■表4: 英語力測定課題の記述統計

	<i>M</i>	95%CI	<i>SD</i>	<i>Min</i>	<i>Max</i>
リーディング	422.81	[410.02, 435.61]	32.35	338	497
リスニング	383.15	[368.78, 397.52]	36.33	316	484
総合点	805.96	[783.24, 828.69]	57.44	668	938

(2) 英作文課題

英作文課題において、Coh-Metrixで算出した指標を受験級ごとに集計した記述統計を表5に示す。調査1の結果を基に、考察をしていく。調査1の結果と同様、Narrativityは級が上がるにつれて、値が減少していく傾向が見られた。この結果により、級が上がるにつれて、質問の内容が難しくなり、馴染みがあまりないトピックになることが示唆された。また、Temporalityに関しては同様の減少傾向が見られたが、模範解答の3級の値とほぼ同じであり、同じような時制や相を使って、

表現していることが示された。

一方、調査1では見られなかった値の上昇傾向がSyntactic SimplicityとVerb Cohesionに見られた。級が上がるにつれて、求められる語数が多くなり、複雑な文構造を使用することができない学習者が似たような馴染みのある文構造を用いて英文を書いている可能性がある。また、同様の傾向がVerb Cohesionにも見られ、求められる語数が多くなるにつれて、同じような動詞を繰り返し使用するようになるためであると考えられる。

■表5: Coh-Metrixによる協力者の英作文の級ごとの特徴

	2級		準2級		3級	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Readability	32.50	8.22	38.50	10.99	35.95	12.04
Narrativity	52.70	22.00	79.51	19.43	81.46	18.70
Syntactic Simplicity	93.98	8.11	89.41	14.82	88.90	25.67
Word Concreteness	30.04	27.44	67.19	31.39	43.72	31.95
Referential Cohesion	50.84	33.12	82.23	28.63	72.58	33.37
Deep Cohesion	79.32	31.06	95.22	20.77	82.90	22.69
Verb Cohesion	77.22	23.09	72.16	30.31	49.49	34.04
Connectivity	5.81	20.14	42.47	35.11	21.29	26.76
Temporality	62.94	31.22	65.10	33.13	68.39	29.76

次に、英作文課題において、Text Inspectorで算出した指標を受験級ごとに集計した記述統計を表6に示す。Percentageにおいて、調査1の模範解答例では、3級が2級と同程度であり、準2級から級が上がるにつれて値が大きくなるという傾向が見られていたが、本研究の結果では、級が上がるにつれて、値が徐々に上がっているという傾向が見られた。Gunning Fogに関しては、上昇傾向が見られており、級が上がるにつれて、

文が長くなったり、2音節以上の単語を使うようになったりしていることが示唆される。Lexical Diversityに関しては、模範解答例で示されたほど数値は高くないが、こちらも増加傾向にあることがわかる。

一方、Metadiscourseに関しては、調査1で見られたような減少傾向は見られず、一貫した結果は得られなかった。

表6: Text Inspectorによる協力者の英作文の級ごとの特徴

	2級		準2級		3級	
	M	SD	M	SD	M	SD
Percentage	45.10	4.38	33.21	8.03	28.85	9.40
Gunning Fog	8.96	1.45	6.12	1.71	5.10	1.60
Metadiscourse	18.67	5.09	23.51	12.11	21.32	3.60
Lexical Diversity	37.03	28.80	12.23	17.22	4.23	10.93

(3) Coh Metrix と Text Inspector で算出された指標と学習者の英語力の関係

分析にあたって、Coh-Metrix と Text Inspector で算出される指標が膨大であるため、算出された受験級ごとの英作文の指標と協力者の英語力をスピアマンの相関分析を行った。協力者の英語力と有意な相関関係が見られたのは、3級の英作文において、Coh-Metrix では、三人称単数代名詞の使用 ($rs = .51, p = .02$)、Text Inspector では、異なり語100語ごとの語の頻度 ($rs = .47, p = 0.3$)、準2級の英作文において、Coh-Metrix では、動詞の重複 ($p = -.45, p = .04$) と 一人称複数代名詞の使用 ($rs = -.424, p = .05$)、Text Inspector では、なし、2級の英作文において、Coh-Metrix では、意図を表す語の使用 ($rs = .50, p = .02$) と 動名詞の利用 ($rs = .46, p = .03$)、Text Inspector では、なしであった。

この結果を基に、それぞれの級において、協力者の英語力(英検IBAの総合点)を従属変数に、有意な相関関係が見られた変数を独立変数に強制投入法による重回帰分析を行った。分析を行う前に、多重共線性の確認を行ったところ、相関係数が .80 を上回る変数はなかった。重回帰分析の結果、それぞれの級に関して、有意な回帰モデルが得られた、 $F(1, 20) = 7.05, p = 0.2; F(2, 19) = 4.76, p = 0.2; F(2, 19) = 4.33, p = 0.3; F(1, 20) = 4.41, p = .05$ 。それぞれの級の結果が表7から表10に示されている。

結果より、3級においては、Coh-Metrix における三人称単数代名詞の使用、英作文で使用される語の頻度が学習者の英語力をそれぞれ26%、18% 予測することができ、準2級においては、動詞の重複と一人称単数代名詞の使用で学習者の英語力を33%、2級においては、意図を表す語と動名

詞の使用で学習者の英語力を31% 予測することができる。ちなみに準2級の英作文においては、 β の値が負であるため、動詞の重複や一人称単数代名詞が少ない協力者ほど、英語力が高いことがわかる。

まとめると、それぞれの受験級において、固有の予測変数は存在するが、共通した予測変数は見つからなかった。したがって、英作文のトピック、もしくは求められる語数によって、英語力を予測する指標は異なる可能性が示唆された。

(4) Coh-Metrix で算出された指標と Text Inspector で算出された英作文得点の関係

(3) と同様、重回帰分析で使用する変数を絞るため、それぞれの受験級ごとの英作文において、Text Inspector で算出された英作文得点 (Percentage) と Coh-Metrix で算出された指標でスピアマンの相関分析を行った。その後、変数間の相関関係を確認し、.80 を超える相関係数の変数は、英作文得点との相関係数が大きい変数を採用した。その結果、重回帰分析の独立変数に選ばれたのは、3級の英作文においては、語(文字数)の長さ ($rs = .47, p = .03$)、Connectivity ($rs = .50, p = .02$)、文間の名詞の重複 ($rs = -.43, p = .04$)、意図を表す動詞の利用 ($rs = -.46, p = .03$)、内容語の語の頻度 ($rs = -.46, p = .03$)、準2級の英作文においては、語(文字数)の長さ ($rs = .51, p = .02$)、内容語の親密度 ($rs = -.62, p = .002$)、2級の英作文において、語数 ($rs = -.70, p < .001$)、語の多様性 ($rs = .53, p = .01$) だった。

この結果を基に、それぞれの級において、協力者の英作文得点 (Percentage) を従属変数に、採用された変数を独立変数に強制投入法による重

■表7: 3級英作文(Coh-Metrix)における強制投入法による重回帰分析の結果

	英語力					
	B	SE B	95%CI	β	t	p
三人称単数代名詞の使用	8.52	3.21	[1.83, 15.21]	0.51	2.65	0.02

CI = confidence interval, $R = .51$, $R^2 = .26$ ($p = .02$)

■表8: 準2級英作文 (Coh-Metrix) における強制投入法による重回帰分析の結果

	英語力					
	B	SE B	95%CI	β	t	p
動詞の重複	-73.38	38.92	[-154.85, 8.09]	-0.35	-1.89	0.08
一人称単数代名詞の使用	-1.91	0.80	[-3.57, -0.25]	-0.45	-2.40	0.03

CI = confidence interval, $R = .58$, $R^2 = .33$ ($p = .02$)

■表9: 2級英作文(Coh-Metrix)における強制投入法による重回帰分析の結果

	英語力					
	B	SE B	95%CI	β	t	p
意図を表す語の使用	0.82	0.43	[-0.09, 1.72]	0.37	1.89	0.07
動名詞の使用	1.43	0.80	[-0.25, 3.12]	0.35	1.78	0.09

CI = confidence interval, $R = .59$, $R^2 = .31$ ($p = .03$)

■表10: 3級英作文(Text Inspector)における強制投入法による重回帰分析の結果

	英語力					
	B	SE B	95%CI	β	t	p
語の頻度	0.00	0.00	[0.00, 0.01]	0.43	2.10	0.05

CI = confidence interval, $R = .43$, $R^2 = .18$ ($p = .03$)

回帰分析を行った。その結果、3級、準2級、2級の英作文に関して、有意な回帰モデルが得られた、 $F(5, 16) = 5.87$, $p = 0.03$; $F(2, 19) = 12.94$, $p < 0.01$; $F(2, 19) = 3.93$, $p = 0.4$ 。それぞれの級の結果は表11から表13に示されている。

表より、有意な予測3級英作文においては、語(文字数)の長さが長く、逆説・付加・比較の接続語、同じ名詞の繰り返し、意図を表す動詞が少なく、内容語の語の頻度が低ければ、英作文の得点が高くなり、準2級英作文においては、語(文字数)の長さが長く、内容語の親密度が低ければ、英作文の得点が高くなり、2級の英作文においては、語数が少なく、語の多様性が低ければ、英作文の得点が高くなることが示された。

それぞれの級に共通する全く同じ予測変数は

見つからなかった。これは英作文のレベルやトピック、書き手の熟達度、利用したルーブリックなどにより英作文の得点に影響を与える指標が異なる(Latifi & Gierl, 2020)という先行研究の結果と一致するものである。一方、英検3級における語の頻度、準2級における語の親密度、2級における語の多様性のような語に関するCoh-Metrixの指標がText Inspectorで算出される英作文の得点に影響を与えていることが示唆された。

■表11: 3級英作文における強制投入法による重回帰分析の結果

	英作文得点					
	B	SE B	95%CI	β	t	p
語(文字数)の長さ	19.36	7.78	[2.82, 35.89]	0.40	2.48	0.03
Connectivity	-0.03	0.09	[-0.15, 0.09]	-0.09	-0.52	0.61
文間の名詞の重複	-10.29	6.44	[-23.94, 3.36]	-0.25	-1.60	0.13
意図を表す動詞の利用	-0.13	0.06	[-0.25, -0.01]	-0.37	-2.30	0.04
内容語の語の頻度	-20.159	7.26	[-35.54, -4.77]	-0.44	-2.78	0.01

CI = confidence interval, $R = .80$, $R^2 = .46$ ($p = .003$)

■表12: 準2級英作文における強制投入法による重回帰分析の結果

	英作文得点					
	B	SE B	95%CI	β	t	p
語(文字数)の長さ	7.14	2.35	[2.22, 12.07]	0.45	3.03	0.01
内容語の親密度	-0.92	0.22	[-1.39, -0.46]	-0.63	-4.19	< 0.01

CI = confidence interval, $R = .76$, $R^2 = .58$ ($p = .001$)

■表13: 2級英作文における強制投入法による重回帰分析の結果

	英作文得点					
	B	SE B	95%CI	β	t	p
語数	-0.10	0.06	[-0.22, 0.03]	-0.58	-1.64	0.12
語の多様性	-1.36	10.97	[-24.31, 21.60]	-0.04	-0.90	0.90

CI = confidence interval, $R = .54$, $R^2 = .29$ ($p = .04$)

4.4 調査2のまとめ

調査2では、Coh-Metrix や Text Inspector で算出された指標の英作文評価への利用可能性を検証するために、日本人英語学習者の英作文に対する Coh-Metrix と Text Inspector の指標を算出し、その指標の英語力の予測率、また、Text Inspector で算出された英作文得点への Coh-Metrix の指標の予測率を調査した。結果として、(1)日本人英語学習者の英作文を英文解析ツールで、分析したところ、受験級が上がるにつれて、英作文の馴染みやすさや時制や相の利用、難易度、語の多様性が上昇傾向にあり、反対に、統語の簡単さや動詞の一貫性には減少傾向が見られた。一方、(2)英作文の受験級(求められるトピックや語数)によって日本人英語学習者の英語力を予測する指標が異なる可能性、また、(3)英文解

析ツールによって算出された英作文得点は、受験級ごと予測変数は異なっていたが、特に、語に関する指標(e.g., 長さ、頻度、親密度、多様性)の影響が大きい可能性が示唆された。

5 結論と今後の課題

本研究では、無料で利用可能なウェブベースの英文解析ツールである Coh-Metrix と Text Inspector の自動英作文評価への利用可能性について2つ調査を行った。まず、調査1では、Coh-Metrix と Text Inspector を用いて、英検が公表している1級から3級の英作文の模範解答例のテキスト分析を行い、テキスト特性を検証した。結果として、受験級が高くなるほど、トピックが

難しくなり、幅広い単語や時制・相を用いた難易度が高い英作文を書くように求められていることがわかった。次に、調査2では、Coh-MetrixとText Inspectorを用いて、実際の日本人英語学習者の英作文のテキスト分析を行い、算出された指標の傾向と英語力との関係、英作文得点との関係を調査した。結果として、英作文のテキスト特性に関しては、受験級が上がるにつれて、英作文の馴染みやすさや時制や相の利用、難易度、語の多様性が増加する傾向にあることがわかり、日本人英語学習者は、級が上がるにつれて内容的にも形式的にも難しい英作文を書くことが示唆された。一方、統語の簡単さや動詞の一貫性は減少傾向にあり、求められる語数が多くなるほど、簡単な馴染みのある文や同じ動詞を繰り返し使うことがわかった。算出された指標と学習者の英語力の関係に関しては、英作文の受験級によって、英語力と関係のある指標が異なることがわかった。求められるトピックや語数が異なることで、英語力を予測する指標が変わることも示唆された。算出された指標と英作文評価との関係性については、英語力との関係性と同様、受験級ごとに英作文評価と関係のある指標が異なっていた。しかし、全ての受験級を通して、長さや頻度、親密度、多様性のような語に関する指標が予測変数に含まれていたことから、語に関する指標が自動英作文評価に影響を与える可能性があることが示唆された。

以上の本研究の結果より得られた教育的示唆としては、以下の点が挙げられる。まず、調査1の結果より、受験級が上がるにつれて、馴染みやすさ、時制・相の利用、語の多様性に一貫した傾向が見られることがわかった。これより、教師が生徒の英作文を評価する際に、語の多様性や時制・相の利用に関して、客観的な指標を用いて、フィードバックすることが可能である。それぞれの受験級で求められている客観的な数値がわかれば、その指標に基づいて、具体的なフィードバックをすることが可能である。例えば、語の多様性に関しては、生徒の英作文に関して、算出された数値が受験級で求められている数値より低ければ、同じ語を使わずに、パラフレーズして書くように指導することが可能であり、リライトした際も、前回の英作文と数値を比較して、どの程度増加した

のか、何が足りないのかを具体的にアドバイスすることが可能である。

また、調査2の結果より、英作文で求められるトピックや語数により、英語力を予測する指標が異なることが示唆された。これより、英文解析ツールを用いて、英作文を評価し、英語力を推定する際は、語数やトピックごとにまず傾向を掴むことが必要である。学習者はトピックや質問に含まれる語や文型を繰り返し使用する可能性がある。したがって、生徒の英語力を推定する際には、トピックや質問に合わせた指標の利用が必要である。さらに、調査2の結果から、英文解析ツールを用いた英作文の自動評価には、語に関する指標が有効である可能性が示唆された。Coh-MetrixやText Inspectorは大規模なコーパスを用いて、語に関する指標を算出しており、信頼性も高い。したがって、生徒の英作文を評価する際には、頻度や多様性といった語に関する指標を評価やフィードバックに有効利用できると考えられる。

本研究は、ウェブベースで利用可能で、読解研究に利用されているCoh-MetrixとText Inspectorを用いた英作文自動評価の妥当性を検証することを目的に行われた。妥当性を十分に検証することはできなかったが、Coh-MetrixやText Inspectorを用いた英作文自動評価の可能性を中心に議論を進めた。一方、本研究には以下のような限界点がある。

まず、調査1では、マテリアルとして英検3級に英作文問題が導入された2017年度から2019年度の英作文解答例を使用した。英作文の解答例は各回に1つしか公表されないため、長文読解テスト等に比べ、分析対象とした英作文の数が少なくなってしまった。さらに、調査2においても、調査を3日間に分けて行ったため、協力者に比べ、分析対象とした英作文の数が少くなってしまった。したがって、今後の研究では、より幅広いマテリアルを分析対象とし、また、協力者の数を増やしたり、協力者の熟達度の幅を広げたりする必要がある。

さらに、各級ごとのトピックが英作文に影響した可能性がある。先述したように、書き手はトピックや質問に含まれている単語や文法を使用する傾向が高い。したがって、今後の研究では、同じトピックで語数が異なる英作文課題や幅広

いトピックで同じ語数の英作文課題を使用したり、結果を一般化する手立てを考えたりする必要がある。

以上のような限界点を考慮し、研究を進めていくことで、英作文自動評価の妥当性についてより説得力のある示唆を出すことができると考えられる。

謝辞

本研究の実施、発表にあたりまして、公益財団法人日本英語検定協会と関係者の皆様、ならび

に選考委員の先生方からの支援をいただき、心より御礼申し上げます。特に、助言者である村木英治先生には、研究の実施および報告書の執筆にあたりご指導いただきましたこと、深く感謝申し上げます。また、筑波大学の卯城祐司先生をはじめ、研究室の先輩・同輩・後輩の皆さんには、本研究の立案から実施、報告書の執筆にあたりまして、ご助言いただきました。心より感謝申し上げます。最後に、本調査に協力してくださった協力者の皆さんに深く御礼申し上げます。

参考文献(*は引用文献)

- * Aryadoust, V. & Liu, S. (2015). Predicting EFL writing ability from levels of mental representation measured by Coh-Metrix: A structural equation modeling study. *Assessing Writing*, 24, 35-58.
- * Bax, S. (2012). Text Inspector, online text analysis tool.
- * Bax, S., Nakatsuha, F., & Waller, D. (2019). Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System*, 83, 79-95.
- * Crossley, S. A., & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In Catrambone, R. and Ohlsson, S. (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp.984-989). Austin, TX.
- * Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life Long Learning*, 21, 170-191.
- * 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, 115-135.
- * Graesser, A. C., McNamara, D.S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36, 193-202.
- * Guo, L., Crossley, S. A., & McNamara, D. S. (2013). Predicting human judgements of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing*, 18, 218-238.
- * 小林雄一郎・金丸敏幸. (2012). 「Coh-Metrix とパターン認識を用いた課題英作文の自動評価」『じんもんこん2012論文集』259-266.
- * Latifi S. & Gierl, M. (2020). Automated scoring of junior and senior high essays using Coh-Metrix features: Implications for large-scale language testing. *Language Testing*.
- * McNamara, D.S., Crossley, S. A., & McCarthy, P. M.
- (2010). Linguistic features of writing quality. *Written Communication*, 27, 57-86.
- * McNamara D. S., Graesser, A. C., McCarthy P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.
- * 文部科学省 (2018). 『学習指導要領（平成30年告示）解説 外国語編 英語編』東京:開隆堂出版.
- * Polio, C. & Friedman, D. A. (2017). *Understanding, evaluating, and conducting second language writing research*. New York: Routledge.
- * 佐野富士子 (2013). 「ライティング」 In JACET (大学英語教育学会) SLA 研究会 (編著), 『第二言語習得と英語科教育法』(pp. 248-261). 東京:開拓社.
- * Shermis, M. D., & Burstein, J. (2003). Introduction. In M. D. Shermis & J. Burstein (Eds.), *Automated essay scoring: A cross-disciplinary perspective* (pp. xiii-xvi). Mahwah, NJ: Lawrence Erlbaum Associates.
- * Zedelius, C. M., Mills, C., & Schooler, J. W. (2019). Beyond subjective judgements: Predicting evaluations of creative writing from computational linguistic features. *Behavior Research Methods*, 51, 879-894.

A 研究部門・報告Ⅲ・英語能力テストに関する研究

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—

研究者:岐阜県／岐阜県立 関高等学校 教諭 福田 陽子(申請時:名古屋大学大学院 在籍)

《研究助言者:寺内 一》

概要

本研究は、ライティングにおける文章の一貫性・結束性に焦点を当て、英語学習者の表現にはどのような言語的特徴があり、それらがエッセイ評価(Essay Score)にどう影響を及ぼしているのかを検証した。

調査1では、日本人英語学習者と英語母語話者によって書かれたのべ1,020エッセイを比較した。その結果、学習者の表現には主語の人称代名詞や、文頭・節間におけるつなぎ語の過剰使用がみられた。また、つなぎ語の誤用によって、内容面における論理の飛躍や因果関係の欠如など、一貫性が損なわれていることも確認された。

調査2では、3名の読み手によって評価されたエッセイ内の一貫性・結束性指標とEssay Scoreとの関連性を分析した。その結果、Scoreと一貫性・結束性には有意な相関があり、特につなぎ語・指示代名詞がエッセイの不自然さを左右する要因になると示された。

高熟達度の学習者ほど、「より自然な表現」ができる、それがエッセイの高評価につながっていることから、つなぎ語・指示代名詞の適切な使用に伴い、内容面の論理性・因果関係を構築する学習を継続的に行なうことが、ライティング能力の習得に貢献する可能性がある。

1

はじめに

グローバル化の急進に伴い、英語教育におけ

るライティングの需要は年々高まっている(Oi, 2002; Daud, & Kassim, 2016)。しかし、英語学習者(以下、学習者)にとってライティングの習得は容易なことではない(Collins, 1998)。文部科学省が実施した日本の高校3年生6万人を対象とした調査結果によれば、CEFRのA2レベル(注1)以上の割合は、「聞く(33.6%)」「読む(33.5%)」「書く(19.7%)」「話す(12.9%)」となり、4技能の中でも産出面であるライティングの習得率が低いと報告されている(『平成29年度英語教育改善のための英語力調査 事業報告』)。その要因の1つとして、ライティングの指導方法が学校教育においてまだ十分に確立されていないことに加え(Oi, 2002; 2008)、ライティングを実践する場が限定的であるため、知識が定着しないと考えられている(Narita, 2018)。こうした状況は日本に限ったことではない。基本的な語彙や文法をマスターしている中・上級学習者であってもライティングにおける不自然・不適切な表現が顕著に表れる(Lorenz, 1998)など、英語を母語としない学習者を対象とした同様の先行研究は複数存在する(Master, 2002; Hopp, Bail, & Jackson, 2020)。こうしたライティングの難しさに直面している中級学習者を研究対象とした場合、語彙や文法以外の問題点として文章の「一貫性」・「結束性」が挙げられる(Carrell, 1982; Collins, 1998)。結束性(cohesion)は文章内の言語的つながり(Crossley, Kyle, & McNamara, 2016)を、一貫性(coherence)は文章内の内容的つながりを

指す (Lee, 2002; Plakans, & Gebril, 2017)。英語の基本ルール (e.g., SVO, SVOC) に従って單語を配列すれば、一見文法的には全く問題が無い文となりうる。しかし、そうした文を並べて文章を作っても、一貫性・結束性に留意していなければ、それは文の「羅列」となってしまう。内容的にまとまりのある文章には、文章内の個々の要素 (例: 単語・句・節・文) の結びつきである「結束性」 (Halliday, & Hasan, 1976) と、内容に矛盾がなく文章の整合性がとれている「一貫性」が不可欠であり (Todd, Thienpermpool, & Keyuravong, 2004), こうした能力は、語彙や文法をマスターするのとは別の能力として重要である。よって本研究ではこの一貫性・結束性に着目し、

1. 学習者・母語話者が書いたエッセイの比較により、学習者の言語的特徴を検証する。

さらに、「読み手」側の視点に立った先行研究が少ないと現状を踏まえ、英語母語話者3名に学習者のエッセイを読んで評価してもらい、

2. そして得られた一貫性・結束性に関わる言語的指標とエッセイ評価 (Essay Score)との関連性を分析する。

以上2点を調査することにより、一貫性・結束性に関わる学習者のライティングの問題点の所在を明らかにする。

(注1) CEFR: "Common European Framework of Reference for Languages" の略で「ヨーロッパ言語共通参照枠」

A2レベル：基礎段階の言語使用者。ごく基本的な個人情報や家族情報、買い物、地元の地理、仕事など、直接的関係がある領域に関しては、文やよく使われる表現が理解できる。簡単で日常的な範囲なら、身近で日常の事柄について、単純で直接的な情報交換に応じることができる。

(各資格・検定試験とCEFRとの対照表 附属資料①参照 文部科学省 平成30年3月 出典はブリティッシュ・カウンシル、ケンブリッジ大学英語検定機構)

(https://www.mext.go.jp/b_menu/houdou/30/03/_icsFiles/afieldfile/2019/01/15/1402610_1.pdf)

2 先行研究

2.1 不自然さの定義

学習者の言語産出には文法面だけでなく、運用面においても「適切ではない表現」 “inappropriate use” が存在するとし、たとえ基本的な英語のルールに従って語彙を配列しても、母語話者から見れば “un-English” とみなされる (Zarifi & Mukundan, 2014)。また、英語の基本をマスターした中・上級学習者であっても、自然な英語表現の欠如が顕著に表れる (Lorenz, 1998)。さらに、EFL 学習者 (English as a foreign language の略) が書いた文章は、語彙・文法ミスが無いとしても、文体がぎこちなく不自然である (Machado, & Tuero, 2013) 等の報告がある。しかし、そもそも適切な英語表現、自然な英語表現の定義が明確ではない。また不自然さを決めるのは誰なのか、読み手であるとすれば判断には個人差が生じることも考えられる。こうした点を踏まえれば、現段階で不自然さを普遍的な定義とみなすことはできないものの、これまでの研究の多くは、文法的な間違いではない・適切とは言えない表現という点で概ね意見は一致している。よって、本研究ではこれらの点を考慮し、不自然さを「文法的な間違いではないが適切とは言えない表現であり、一般的な母語話者には見られない学習者特有の言語表現」と定義した。

2.2 文章の結束性

Halliday & Hasan (1976) は結束性を “tie” (p. 3) と表現し、文章の語彙的・文法的つながりを示す「結び・つなぎ」の働きであると述べている。結束性は例えば人称代名詞 (例:I, he, she, you, we, they), 指示代名詞 (例:this, that), 接続詞 (例:and, but), 副詞 (句) (例:however, therefore, for example) のように明示的であると言われ、これらは全て結束性のつなぎとして機能している (Halliday & Hasan, 1976)。本研究では接続詞・副詞 (句) を総称してつなぎ語と表記する。例えば、

1. (a) Tom is my brother. Tom is kind.
(結束性なし)
○(b) Tom is my brother. He is kind.
(結束性あり)

1. (a) は2つのTomが別人物を指す可能性があり、文章間に結束性が無く意味的つながりも存在しない。よって、文章に一貫性が無いと考えられる。一方、(b)はHe=Tomを指し、両者は同一人物を表すことから、文章間に結束性が生まれる。Heはつなぎの働きを持つ“tie”として機能し、一貫性のある文章だと言える。(a)(b)はどちらも文法的間違いは存在しないが、(a)は結束性という点で不自然な表現である。次の例は、

2. (a) Tom was very uncomfortable,
and he fell asleep. (結束性なし)
○(b) Tom was very uncomfortable,
but he fell asleep. (結束性あり)

2. (a) はandで2つの節を結んでいるが、前後の節は意味的繋がりが不自然で結束性・一貫性の無

い文章である。一方、(b)はbutで2つの節をつなぎ、内容的に矛盾が無く文章に一貫性があると言える。このようにつなぎ語によって、文章の一貫性・結束性は強化されたり、損なわれたりすると考えられる。

また、結束性は「情報構造」とも密接な関係にある(Halliday & Hasan, 1976)。情報構造は、情報の伝わりやすさを意識した基本的な語彙配列のルールである。これは読み手にとって前出している旧知の情報を「旧情報」、前出していない新しい情報を「新情報」と呼び、旧情報⇒新情報の順で語彙を配列する。旧情報に新しい情報を追加していくことで、読み手・聞き手に情報が伝わり易くなるというものである(Bloor & Bloor, 2013)。例えば、下段の3. (a)はTomの話題から、2文目で突如Mary will get married toが配置され、旧情報であるhimが文末にきている。情報が新情報⇒旧情報へと基本とは逆に流れており、読んだ際に情報の伝わりにくさ・文章のぎこちなさを感じる。一方、(b)は2文目He=1文目Tomを指し、同様に3文目That=前文を指すことから旧情報⇒新情報へと情報が基本に沿って追加されており、内容がスムーズに理解できる。つなぎの働きであるhe, him, thatをどこに配置するかによって結束性は変化し、それが文章の一貫性にも影響を及ぼすと考えられる。例文1.~3.はHalliday & Hasan (1976), Bloor & Bloor (2013)を参考にして本研究者が自作したものである。

3. (a) Tom is my brother. Mary will get married to him. We are happy to hear that (news). (結束性なし)
- | | | | |
|-----|-----|-----|-----|
| 新情報 | 旧情報 | 新情報 | 旧情報 |
|-----|-----|-----|-----|
- (b) Tom is my brother. He will get married to Mary. That makes us happy. (結束性あり)
- | | | | |
|-----|-----|-----|-----|
| 旧情報 | 新情報 | 旧情報 | 新情報 |
|-----|-----|-----|-----|

2.3 文章の一貫性

一貫性は、内容に矛盾がなく、文章が首尾一貫しているか、といった文章内の内容的つながりを

指す(Carrell, 1982)。しかし、2.2で示したように言語表現から判断ができる結束性とは異なり、一貫性は暗示的で、文章を読まなければ確認することできない(Plakans, & Gebril, 2017)。また、

読み手の背景知識量や主観的判断も影響すると言われており(Todd, Khongput, & Darasawang, 2007), 文章に一貫性があるかどうかを客観的に測定するのは容易ではない。Farghal (1992) は “the cohesive harmony should be a continuum - the higher the cohesive harmony, the more coherent the text is, and vice versa, the lower the cohesive harmony, the less coherent the text is.” (p. 45) と述べ, 結束性が上手く機能していれば, 文章内の一貫性がより促進され, 逆もまた然りだと報告しており, 一貫性・結束性は密接な関係にあると考えられる(Grabe & Kaplan, 1996)。

以下に一貫性・結束性が不自然な文章の一例として日本人大学生が書いたエッセイの一部抜粋を記す。本研究で使用した縦断的学習者コーパス「NICEST」(3.3参照)に収められているエッセイから一部を抽出したものであり, トピックに対して賛成の立場で意見を述べている。

これは文間(文と文の間)・節間(節と節の間)のつなぎ目に I know, I think が繰りかえし使用された不自然な文章である。書き手である I の繰り返しの使用は結束性が無い(Halliday & Hasan, 1976)と考えられており, それによって一貫性も損なわれ, 筆者の伝えたい情報が読み手に適切に伝わらない参考例である。

トピック: Do you agree or disagree with the following statement?

Successful people try new things and take risks rather than only doing what they already know how to do well. Use reasons and examples to support your answer.

エッセイの抜粋:

I agree with that given idea.

I know that some people work on the job which his father or grandfather develops and

I also know they have lots of money, but I don't think they succeed on the job.

I think they are only people that can change to other people.

I think successful people are the people who can never change to others and

I also think they try new things and take risks rather than only doing what they already know how to do well.

2.4 自然さの重要性

なぜ不自然な表現ではいけないのか。Pinker (2015) は, 理解し難い文章が世界には溢れていると述べ, さらに伝えたい情報が読み手に正しく伝わっていないことに書き手が気付いていない, と警笛を鳴らしている。ライティングもスピーキングと同様にコミュニケーションの手段であり, 書き手が書きたいことを書けば良い訳ではなく, 読み手に情報を伝えることを意識する必要がある。

実際スピーキング研究では学習者特有の発音やアクセントがコミュニケーションを阻む要因になるとの報告があり(Luo, Gu, & Wang, 2016), 学習者は自然に近づく発音やアクセントを追及している。同様にライティングも, 不自然・不適切な表現が円滑なコミュニケーションを阻む要因になり得るが, スピーキングとは異なり対話者が目の前に存在しないというライティングの特性があるため, 読み手を想定し情報を適切に伝えようとする意識が薄いと言える。実際, 読み手側

の視点に立った先行研究も少ないとから、学習者の表現がどのように読み手に伝わり評価されるのかを客観的に検証することで、本研究では自然さの重要性を捉えていきたい。

2.5 一貫性・結束性の不自然さに関する先行研究

一貫性・結束性の不自然さに関する先行研究は、例えばEFL学習者のつなぎ語の過剰使用(Ersanli, 2015)や、日本人学習者のand, I, I think, weの過剰使用(Natsukari, 2012; Kojima, 2019)など、コーパスを使用し、母語話者との比較によって学習者特有の表現を報告している。こうした研究に加え、学習者の表現がエッセイ評価と関連があるのかを調査する試みが2000年代は急増する。これはWeb上の言語指標算出ツール、例えば、Coh-Metrix(Graesser, McNamara, Louwerse, & Cai, 2004)等の開発により、エッセイ内の言語的指標をコンピューターで瞬時に算出が可能となったためであり、こうして一貫性・結束性とエッセイ評価との関連性を調査する研究が加速したと考えられる。

2.6 「一貫性」とエッセイ評価との関連性

Crossley & McNamara(2010)はEFL大学生が書いた184エッセイの語彙・文法・文章構成・一貫性・結束性等の言語的指標と評価者によるエッセイ評価との関連性を調査した。また、Plakans, & Gebril(2017)はTOEFL iBTのエッセイを使用し、3つの言語的指標(文章構成・一貫性・結束性)と評価者によるエッセイ評価との関連を調査した。これらの研究結果は、エッセイ評価と一貫性には関連性があると結論づけている。同様の研究結果が複数存在することから、一貫性がエッセイ評価に影響を及ぼすことは概ね一致した見解である。一方、結束性とエッセイ評価との関連性は意見が二分している。

2.7 「結束性」とエッセイ評価との関連性

Liu & Braine(2005)はEFL大学生が書いた50のエッセイを使用し、つなぎ語の使用頻度とエッセイ評価との関連性を分析し、つな

ぎ語の使用頻度が高ければエッセイ評価が上がる調査結果を報告している。Yang & Sun(2012)は中國人大学生のつなぎ語の正用・誤用を調査し、正用の頻度が高ければエッセイ評価が上がると報告している。これらの先行研究は、結束性とエッセイ評価との関連性がプラス(positive)の効果であると提唱している。一方、Crossley, Weston, Sullivan, & McNamara(2011)は高校生・大学生が書いた194のエッセイを調査し、つなぎ語の使用頻度が低い程、より熟達した書き手であると結論づけている。さらに、McNamara, Crossley, & Roscoe(2013)も初級学習者ほどつなぎ語を多く使用する傾向があると指摘している。これらの先行研究は結束性とエッセイ評価との関連性がマイナス(negative)に機能していることを示している。こうした結果から、結束性とエッセイ評価にプラス、マイナスどちらの関連性があるかは研究間で意見が一致していない。

2.8 先行研究の問題点と本研究の課題

これまで本研究における不自然さを定義し、文章内における一貫性・結束性の働きと、それらが不自然であるとどのような問題が起きるのかを整理してきた。関連する先行研究から、学習者にはつなぎ語や人称代名詞の過剰使用があること、また、一貫性がエッセイ評価に影響を及ぼしていることは研究間で概ね同意が得られている。しかし、結束性とエッセイ評価との関連性は研究間で意見が分かれ、結果が一致していないことが確認できた。よって、結束性により重点を置き、調査1では母語話者との比較によって学習者の言語的特徴が何であるかを検証する。調査2では、読み手によって評価された一貫性・結束性指標とエッセイ評価との関連性を分析し、結束性に影響を及ぼしている具体的な指標を検証する。さらに一貫性に関して、読み手にアンケート・インタビュー調査を行い、こうした統計的な量的分析・質的分析の両アプローチにより本研究を実施する。

3 調査1

3.1 目的

調査1の目的は、日本人英語学習者・英語母語話者によって書かれたエッセイの比較によって、結束性に関わる学習者の言語的特徴を明らかにすることである。検証課題は次の2点である。

課題1-1

母語話者には見られない学習者特有の言語的特徴は何か。

課題1-2

学習者にはつなぎ語や人称代名詞の過剰使用といった先行研究と同様の特徴が見られるか。

3.2 使用するデータ

本研究は、名古屋大学で構築された縦断的学習者コード「NICEST」(科学研究費補助金・基盤研究(B)16H03444, 研究代表者: 杉浦正利, <https://ncest-sugiura.blogspot.com>) のうち、日本人学習者(以下, L2)が書いた860エッセイと、母語話者(以下, L1)が書いた160エッセイを含む、のべ1020のエッセイを使用した。

3.3 研究対象者

3.3.1 L2学習者

860エッセイを執筆したL2学習者は、日本語を母語とする112名の大学生であった。平均年齢は19.1歳(年齢幅18-23歳)で、大学での専攻は工学・理学・医学であった。2名を除く110名のほとんどが海外滞在歴も無く、日本の学校教育における英語学習を11~12歳から開始している。自己申告のあった59名のTOEFL ITPの平均点は501.8点(得点幅427-606点)であることから英語の熟達度は中級レベルCEFR B1(注2)と推定された。申告者59名からは点数開示の承諾が得られている。

(注2) B1レベル: 自立した言語使用者。仕事、学校、娯楽などで普段出会うような身近な話題に

ついて、標準的な話し方であれば、主要な点を理解できる。その言葉が話されている地域にいるときに起こりそうな、たいていの事態に対処することができる。身近な話題や個人的に関心のある話題について、筋の通った簡単な文章を作ることができる。(引用・出典は注1と同じ)

3.3.2 L1協力者

160エッセイを執筆したL1協力者は20名の英語母語話者であった。平均年齢は25.6歳(年齢幅18-34歳)で、国籍は米国・英国・カナダであった。職業は英語教師・エディターなど大学卒業資格を保持した日本在住の英語母語話者である。

3.4 エッセイの収集方法

L1・L2が執筆したエッセイは、環境・教育・福祉など毎回異なったテーマに関して、辞書を使わせず30分間で自分の意見を述べるpersuasive essayである。L1は一日で8つのエッセイを、L2は週一回ずつ8週間に渡り8つのエッセイを執筆した。パソコンで執筆したのち、全てのエッセイはCriterionのETSエッセイ自動採点システム(The Criterion® Online Writing Evaluation service from ETS. <https://criterion.ets.org/criterion/default.aspx>)のGrade12の評価基準により1~6段階評価(以下, Essay Score)が付けられた。エッセイ添削において、専門家とCriterionとを比較した先行研究によれば、Criterionの精度は専門家に劣らず信頼性が高いと報告があるため(Burstein & Wolska, 2003), このEssay Scoreを調査2の分析で使用した。

以下、提示されたトピックの1例である。

Do you agree or disagree with the following statement?
The best way to travel is in a group led by a tour guide.
Use reasons and examples to support your answer.

3.5 L1・L2エッセイの基本情報

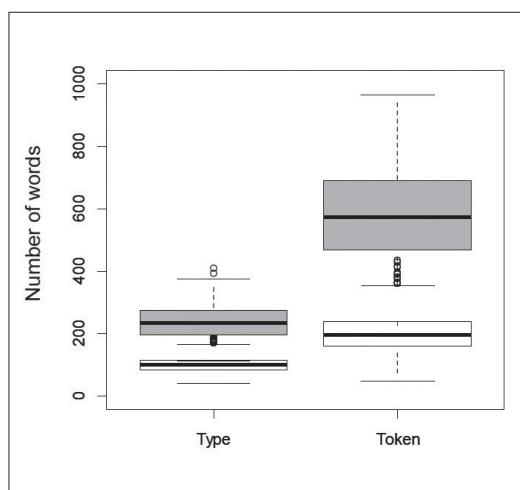
表1はL1・L2エッセイの基本情報として、Essay Score, Type, Token(表1の注)の平均値を比較したものである。表1によれば、Essay Scoreの平均はL1が5.21, L2が2.95であり、両者には大きな開きがある。また、Typeの平均語数

はL1が234.9, L2が102.2、さらに、TokenはL1が572.6, L2が203.4とどちらも2倍以上の開きがある。図1,2の比較図や散布図からも明らかのように、L1・L2がトピック・制限時間を揃え、同一条件下でエッセイを執筆した場合、両者の言語産出能力の違いは歴然としている。

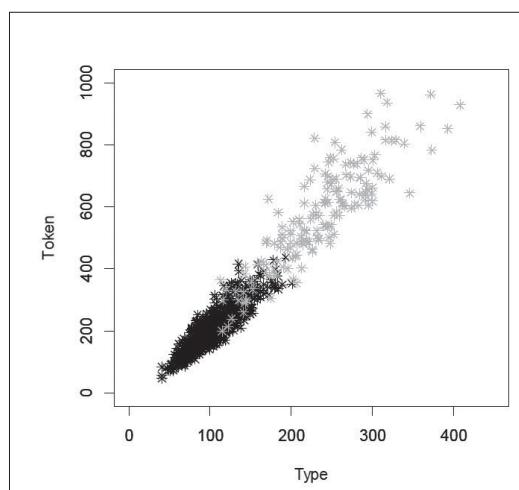
■表1: L1 vs. L2 Essay Score, Type, Tokenの平均値の比較

エッセイ執筆者	Essay Score	Type	Token
L1英語母語話者(n=160)	5.21	234.9語	572.6語
L2日本人学習者(n=860)	2.95	102.2語	203.4語

(注)Essay Score: CriterionのETSエッセイ自動採点システムによって評価されたエッセイスコア、数値が高いほどエッセイの評価が高い(3.4参照)
Type: 1つのエッセイ内の異語数、Token: 1つのエッセイ内の総語数



■図1:L1(グレー)vs.L2(白)のType, Tokenの比較図



■図2:L1(グレー)vs.L2(黒)のType, Tokenの散布図

3.6 L2エッセイの基本情報

より詳細な検証を行うため、L2の860エッセイをEssay Scoreに基づいて下位(Score1, 2),

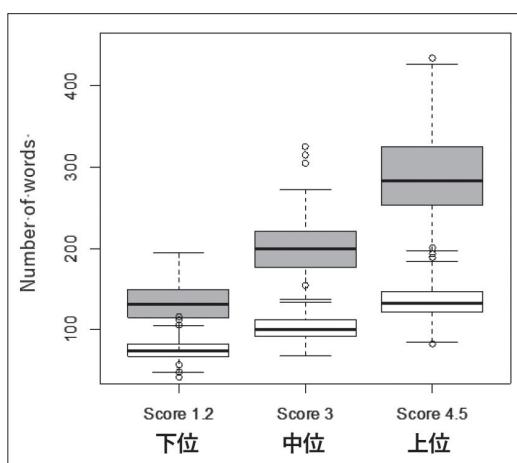
中位(Score3), 上位(Score4, 5)の3つのレベルに分類した。表2によれば、L2の下位、中位、上位、L1の順でType, Tokenの平均語数が大きく増加していることが分かる。また、1エッ

■表2: レベル別のType, Token, 文数, 文長の平均値の比較

レベル	エッセイ数	Essay Score	Type	Token	平均文数	平均文長
L1	160	3-6	234.9	572.6	23.3	27.1
L2上位	175	4-5	134.9	290.6	18.2	16.4
L2中位	490	3	101.4	200.9	13.7	15.8
L2下位	195	1-2	75.0	131.2	10.1	13.7

(注)平均文長:一文当たりの平均単語数

セイ内の平均文数、平均文長(表2の注)もレベルが上がるにつれて増加を示しておりEssay Scoreと比例した関係性であると言え、これらは図3によっても確認ができる。先行研究によりType, Tokenと熟達度の関係性は比例するとの報告があるため(Engber, 1995; Watkins, Kelly, Harbers, & Hollis, 1995), 本研究ではこの3段階のレベルを熟達度とみなした。



■図3:L2のレベル別Type(白),Token(グレー)の比較

3.7 抽出手順

調査1の言語的指標の抽出手順は以下の通りである。

1. コーパス内の1,020エッセイを読み、抽出対象である言語的指標を目視で確認する。
2. 言語的指標を文頭・節間・主語に分類し、CHAT (Codes for the Human Analysis of Transcripts. <https://talkbank.org/manuals/>

CHAT.pdf)形式で分析コードを入力する。

3. 統計分析フリーソフトR version 3.5.2 (R Core team, 2019)を用いて文頭・節間・主語別に指標を抽出し、それぞれの使用頻度を調べる。具体的な方法は3.8で記す。

3.8 抽出方法

抽出方法はHalliday & Hasan(1976)の結束性の分類に基づき、4点の言語的指標を抽出した。

1. 文頭のつなぎ語(例: first, however, in conclusionなど文頭・主語の前にあるもの)
2. 主語(例:I, you, she, it, peopleなど)
3. 節間のつなぎ語(例:and, butなど文中の節と節を結ぶもの)
4. I think, I believe(例:I think + 節, I believe + 節の形式であるもの)

図4に、エッセイデータと入力した分析コードの参考例を示す。

例えば6行目、*JAN0066は日本人のデータであることを意味し、エッセイは下記のように1行1文ずつ収められている。その下に,% cod: howeverと文頭つなぎ語の分析コードを入力し、いずれも存在しない場合は0、また複数存在する場合はコードを追加して入力していく。主語(% sub:)等も同様の手順である。こうしてCHAT形式で分析コードを書き入れ、それをもとに分析タグの抽出と頻度計算作業を実施した。

```
*JAN0066: Nowadays , cars are the most major way to go to another place and some peple say that more and more people will
%cod: nowadays
%sub: cars
%co1: and
%co2: 0
*JAN0066: However I strongly believe that there willl be fewer cars in use than there are today for the following reasons
%cod: however
%cod: strongly
%sub: there
%co1: 0
%co2: I believe
%P
*JAN0066: First of all , cars are bad for the environment .
%cod: first of all
%sub: cars
%co1: 0
```

■図4:エッセイデータと入力した分析コードの例

3.9 検証方法

各指標の使用頻度率の算出には、2通りの方法を採用した。

検証1：

1エッセイあたりで算出し、L1・L2を比較

検証2：

200語あたりで算出し、L1・L2を比較

「検証1」は、1エッセイあたりの使用頻度率を「出現回数÷総エッセイ数」でレベル別に計算した。具体的な算出方法は表3の(注)に記す。これは1つのエッセイ内にどのような表現が頻繁に見られるか、というエッセイ内の特徴を知る上で有効な手段である。しかし、3.5, 3.6で述べたように1エッセイあたりのTokenが各レベルによって大きく異なるため、この方法では使用頻度率を単純に比較することができないという弱点がある。そのため「検証2」は、200語あたりの使用頻度率を「出現回数÷総語数×200語」で計算した。具体的な算出方法は表4の(注)に記す。200語という数字は、L2の1エッセイあたりの平均Token(203.4語)を参照した(表1)。2通りの方法で、より多面的な角度からL1・L2の比較検証を試みた。

次の調査1の結果と考察では、主にL2中位とL1での比較を記している。L2中位のエッセイ数は490/860と全体の過半数を超えており、L2の特徴を最も表していると考えられたためである。

4

調査1の結果と考察

4.1 文頭つなぎ語

文頭つなぎ語の頻出語上位7項目は、second, so, in conclusion, for example, first, however, butである。L2中位のfirst, second, in conclusionの1エッセイあたりの使用頻度率は0.46, 0.59, 0.50であるのに対し、L1は0.04, 0.03, 0.21である(表3)。両者の違いは歴然としており、L1のエッセイにはfirst, second, in conclusionという文頭表現はほとんど見られず、特にfirst, secondの使用は皆無と言ってもよい。一方、L2エッセイはfirst, secondにfirstly, secondlyも含めれば全体の6~7割程度がこの順序付けを用いてエッセイを執筆しており、L2学習者の特徴と言える。表3・表4には存在しないが、L1はfirst, secondの代わりにthen(1エッセイあたり0.63), also(0.57)を用いた表現が多く、一方、L2にはそれらの使用はほとんど見られなかった。

図5・図6は表3・表4を折れ線グラフにしたものである。図5はL1を基準にして1エッセイあたりの使用頻度率の高いものから順に並べており、図6は図5を200語あたりに変換したグラフである。図5によれば、L1のA. however, C. butの使用頻度が高いことから、howeverなど逆接的な表現を多く採用してL1がエッセイを構築していることが分かる。一方、図6を見ると、ほぼ全ての指標においてL2の使用頻度率が高く、L2は文頭つなぎ語の過剰使用であると総合的に判断することができ、これらは先行研究を支持する結果である(e.g., Ersanli, 2015)。また、日本では文頭butの使用は推奨されていないが、L1のエッセイには文頭のbutの使用も見受けられた。また、図6をレベル別に見ると、L2下位、中位、上位と高熟達度になるほど、文頭つなぎ語の使用傾向がL1に近づいていくことが分かった。first, second, for exampleなど文頭つなぎ語の使用は学習者の学校教育での指導が影響していると考えられる。

■表3: 文頭つなぎ語の使用頻度Top 7(1エッセイあたり)

頻出語 Top7	L2 上位 (n=175)		L2 中位 (n=490)		L2 下位 (n=195)		L1 (n=160)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
second	95	0.54	287	0.59	74	0.38	4	0.03
so	86	0.49	267	0.54	102	0.52	29	0.18
in conclusion	95	0.54	246	0.50	48	0.25	33	0.21
for example	104	0.59	228	0.47	51	0.26	67	0.42
first	75	0.43	226	0.46	66	0.34	7	0.04
however	115	0.66	225	0.46	36	0.18	125	0.78
but	30	0.17	115	0.23	61	0.31	67	0.42

(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率

1エッセイあたりの使用頻度率の計算方法: 出現数÷レベル別総エッセイ数

例えば、L2 上位の second の場合は、95÷175=0.54となる

■表4: 文頭つなぎ語の使用頻度Top 7(200語あたり)

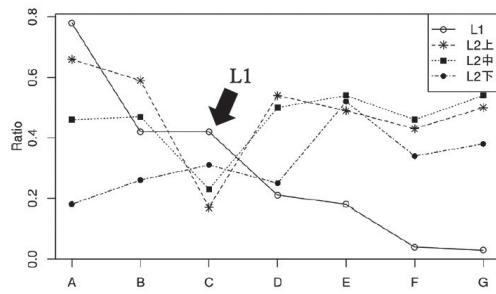
頻度語 Top7	L2 上位 (n=50869)		L2 中位 (n=98472)		L2 下位 (n=25589)		L1 (n=91614)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
second	95	0.37	287	0.58	74	0.58	4	0.009
so	86	0.34	267	0.54	102	0.80	29	0.06
in conclusion	95	0.37	246	0.50	48	0.38	33	0.07
for example	104	0.41	228	0.46	51	0.40	67	0.15
first	75	0.29	226	0.46	66	0.52	7	0.015
however	115	0.45	225	0.46	36	0.28	125	0.27
but	30	0.11	115	0.23	61	0.47	67	0.15

(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率

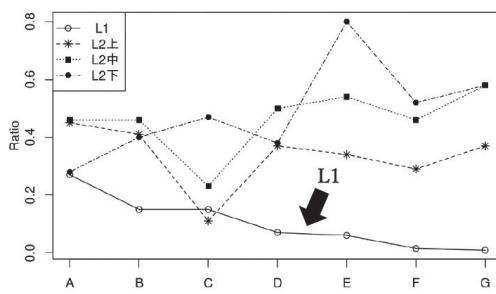
200語あたりの使用頻度率の計算方法: 出現数÷レベル別総語数×200

例えば、L2 上位の second の場合は、95÷50869(175エッセイの総語数)×200=0.37となる

文頭つなぎ語 使用頻度Top7



■図5:1エッセイあたりの使用頻度率



■図6:200語あたりの使用頻度率

A:however B:for example C:but D:in conclusion E:so F:first G:second

4.2 節間のつなぎ語

節間のつなぎ語の頻出語上位6項目は、and, but, because, so, however, asである。1エッセイあたりのand, but, asのL2中位の使用頻度率は0.91, 0.44, 0.02であり、一方、L1は2.48, 1.36, 0.76であることから、L1の使用の多さが目立つ(表5、図7)。しかし、これは1エッセイあたりのToken平均がL1は572.6であるのに対し、L2中位は200.9と2倍以上の差があるためであり(表2)、L1の使用頻度率が高いことは理解ができる。一方、表6、図8の200語あたりで見てみれば、and, butの使用はL1・L2で類似が見られるが、L1の平均文

長が27.1に対して、L2中位が15.8語であることを考えれば(表2)、L2が短文の中でand, butを繰り返し使用している可能性もあり、文章の不自然さの要因になり得るであろう。L1はasを使った表現を好む一方、L2はbecauseやsoの過剰使用が特徴的である(図8)。L2はand, but, because, soを初期段階で学習するため使用に慣れているのか、または語彙力が乏しく多様な表現が出来ないため、同じ語彙の使用を繰り返してしまうのではないかと推測される。また、質的分析の中で、節間のつなぎ語の誤用がしばしば確認されたが、その点に関しては後ほど6.4で取り上げる。

■表5: 節間のつなぎ語の使用頻度Top 6(1エッセイあたり)

頻度語 Top6	L2 上位 (n=175)		L2 中位 (n=490)		L2 下位 (n=195)		L1 (n=160)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
and	253	1.45	446	0.91	89	0.46	396	2.48
but	118	0.67	218	0.44	52	0.27	217	1.36
because	137	0.78	263	0.54	63	0.32	103	0.64
so	85	0.49	203	0.41	47	0.24	65	0.41
however	5	0.03	24	0.05	0	0.00	45	0.28
as	16	0.09	10	0.02	5	0.03	122	0.76

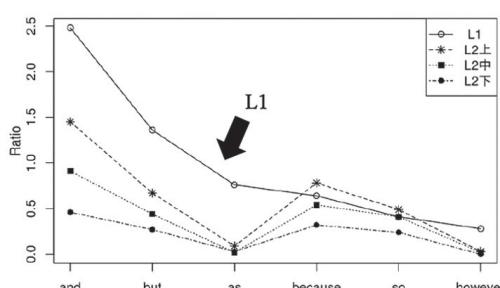
(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率

■表6: 節間のつなぎ語の使用頻度Top 6(200語あたり)

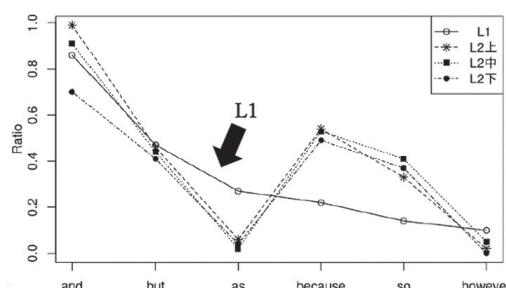
頻度語 Top6	L2 上位 (n=50869)		L2 中位 (n=98472)		L2 下位 (n=25589)		L1 (n=91614)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
and	253	0.99	446	0.91	89	0.70	396	0.86
but	118	0.46	218	0.44	52	0.41	217	0.47
because	137	0.54	263	0.53	63	0.49	103	0.22
so	85	0.33	203	0.41	47	0.37	65	0.14
however	5	0.02	24	0.05	0	0.00	45	0.10
as	16	0.06	10	0.02	5	0.04	122	0.27

(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率

節間のつなぎ語 使用頻度Top6



■図7:1エッセイあたりの使用頻度率



■図8:200語あたりの使用頻度率

4.3 主語

主語の頻出語上位7項目は、I, it, you, we, they, there, thisである。I, you, weの、L2中位の使用頻度率は2.27, 1.16, 1.07、一方、L1は2.11, 1.05, 0.84であることから(表7)、1エッセイあたりでは、L1も人称代名詞の使用が多いことが分かる。しかし、200語あたりに換算すれば、L2は2.25, 1.15, 1.06であるのに対し、L1は0.74, 0.37, 0.29であることから(表8)、L2の人称代名詞I, you, weの過剰使用は明らかである。この結果

は図9,10の比較によっても確認することができ、L2の人称代名詞の過剰使用は、概ね先行研究を支持する結果である(e.g., Yang, & Sun, 2012)。また、thisの使用に関して、L2中位は0.29であるのに対し、L1は1.25と大きな開きがある(表7)。主語における指示代名詞thisの使用は、L1が情報構造(2.2参照)を意識して執筆している可能性があることを示唆している。L1のthisの使用に関しては6.4.3で参考例を示す。

■表7: 主語の使用頻度Top 7(1エッセイあたり)

頻度語 Top7	L2上位(n=175)		L2中位(n=490)		L2下位(n=195)		L1(n=160)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
I	451	2.58	1110	2.27	283	1.45	337	2.11
it	287	1.64	731	1.49	171	0.88	442	2.76
you	275	1.57	566	1.16	178	0.91	168	1.05
we	242	1.38	523	1.07	177	0.91	135	0.84
they	174	0.99	422	0.86	100	0.51	197	1.23
there	133	0.76	376	0.77	124	0.64	184	1.15
this	129	0.74	142	0.29	47	0.24	200	1.25

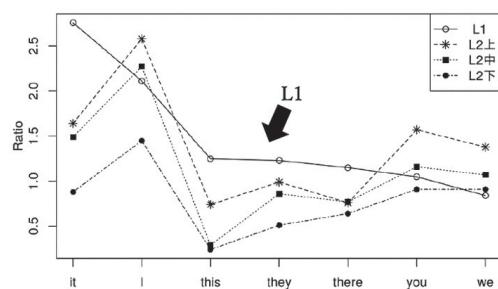
(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率

■表8: 主語の使用頻度Top 7(200語あたり)

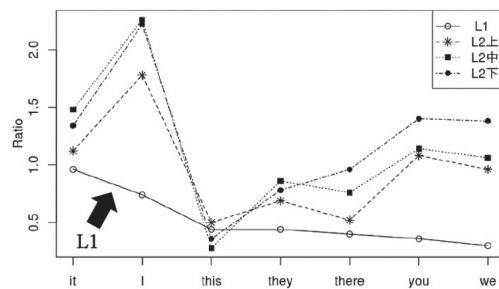
頻度語 Top7	L2上位(n=50869)		L2中位(n=98472)		L2下位(n=25589)		L1(n=91614)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
I	451	1.77	1110	2.25	283	2.21	337	0.74
it	287	1.13	731	1.48	171	1.34	442	0.96
you	275	1.08	566	1.15	178	1.39	168	0.37
we	242	0.95	523	1.06	177	1.38	135	0.29
they	174	0.68	422	0.86	100	0.78	197	0.43
there	133	0.52	376	0.76	124	0.97	184	0.40
this	129	0.51	142	0.29	47	0.37	200	0.44

(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率

主語の使用頻度Top7



■図9:1エッセイあたりの使用頻度率



■図10:200語あたりの使用頻度率

4.4 I think, I believe

1エッセイあたりのI thinkの使用頻度率に関して、L2中位は0.77であるのに対し、L1は0.58であることから、L1もI thinkをしばしば使用していることが確認された(表9、図11)。しかし、200語あたりで見れば、L2は0.77であるのに対し、L1は0.20であることから、学習者のI thinkの過剰使用は明らかであり(表10、図12)，これらは先行研究を裏付ける結果である(e.g., Natsukari, 2012)。一方、1エッセイあたりのI believeの使

用頻度率は、L2中位が0.12、L1は0.69とL1が非常に高く(表9、図11)，I believeの使用はL1エッセイの特徴だと言える。

■表9: I think等の使用頻度Top 4(1エッセイあたり)

頻度語 Top4	L2上位(n=175)		L2中位(n=490)		L2下位(n=195)		L1(n=160)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
I think	149	0.85	378	0.77	141	0.72	92	0.58
I believe	39	0.22	57	0.12	14	0.07	110	0.69
I do not think	11	0.06	27	0.06	18	0.09	13	0.08
I know	8	0.05	14	0.03	1	0.01	10	0.06

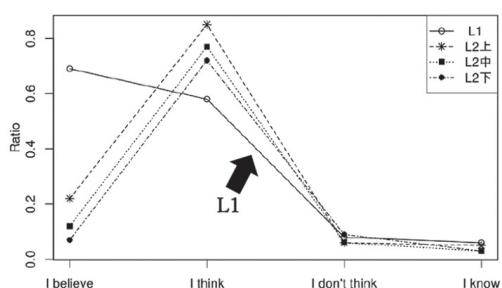
(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率, I do not thinkにはI don't thinkも含む

■表10: I think等の使用頻度Top 4(200語あたり)

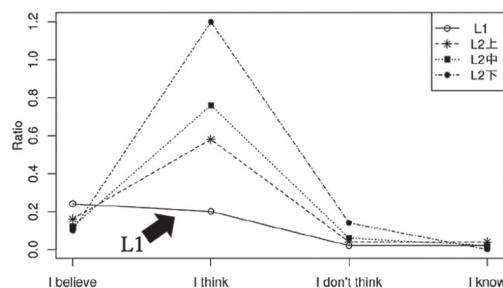
頻度語 Top4	L2上位(n=50869)		L2中位(n=98472)		L2下位(n=25589)		L1(n=91614)	
	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio	Freq.	Ratio
I think	149	0.59	378	0.77	141	1.10	92	0.20
I believe	39	0.15	57	0.12	14	0.11	110	0.24
I do not think	11	0.04	27	0.05	18	0.14	13	0.03
I know	8	0.03	14	0.03	1	0.01	10	0.02

(注) Freq.=使用(出現)頻度数, Ratio=使用(出現)頻度率, I do not thinkにはI don't thinkも含む

I think, I believeの使用頻度



■図11:1エッセイあたりの使用頻度率



■図12:200語あたりの使用頻度率

4.5 調査1のまとめ

課題1-1は、母語話者には見られない学習者特有の言語的特徴は何か、であった。結果から、学習者は結束性に関わるほとんどの指標において全体的に過剰使用の傾向が見られた。特に、文頭のつなぎ語は、first, second, in conclusion, さらに, so, for example が過剰に使用されていた。節間のつなぎ語は、because, so, また主語における人称代名詞は I think を含め I, you, we の過剰使用が特徴的であった。一方、不足に関しては、文頭つなぎ語の then, also, 節間のつなぎ語 as, 主語の this であり、これらが適切に使用されていないことが明らかとなった。

課題1-2は、学習者にはつなぎ語や人称代名詞の過剰使用といった先行研究と同様の特徴が見られるか、であった。課題1-1の結果より、先行研究と概ね同様の結果が得られた。特に「文頭つなぎ語」の過剰使用は、先行研究でもまだ十分に調査が進んでおらず、新たな発見となった。

5 調査2

5.1 目的

調査2の目的は、3名の読み手(以下、評価者と記す)によって評価された一貫性・結束性指標と、エッセイ自動採点システムによって評価された Essay Score (3.4参照)との関連性を分析することである。検証課題は次の3点である。

課題2-1

一貫性・結束性と Essay Score に関する連性はあるか。

課題2-2

2-1の結果が得られた場合、結束性に影響を及ぼしている具体的な指標は何か。

課題2-3

評価者へのアンケート・インタビュー調査から学習者が抱えている一貫性・結束性の問題点は何か。

5.2 3名の評価者

エッセイの評価を依頼したのは、米国出身の3名の英語母語話者である。先行研究(e.g., Crossley, & Mcnamara, 2010)では、英語教育を専門とし、かつ英作文の指導経験が3年以上ある母語話者を評価者に選定しているため、本研究でも先行研究と同様の経歴を持ち、言語学的背景知識が十分にあると判断した母語話者に評価を依頼した(表11)。

■表11: 3名の評価者情報

評価者1:米国出身	30代男性 英語教育を専門とする大学院生
評価者2:米国出身	50代男性 中学生を対象とした塾講師
評価者3:米国出身	30代男性 高校生を対象とした英語指導助手

5.3 データの収集方法

3名の評価者には同一の100エッセイを読み、評価シートを用いて一貫性・結束性指標を5段階で評価するように依頼し、のべ300データ(100エッセイ×評価者3名分)を収集した。対象とした100エッセイは、調査1で用いた1,020エッセイの中から抽出されたものである。詳細は以下5.4, 5.5で記す。

5.4 調査対象とした100エッセイの内訳

調査対象とした100エッセイは、レベル別での比較を行うためL2下位・中位・上位から90エッセイ、L1から10エッセイとした。特に、L2の特徴を最も表していると推測されるスコア3,4を多く選定し、一方エッセイの基本的スタイルを十分に満たしていないスコア1は対象から除外した。データに偏りがないようレベル別に分類した上で、100エッセイをランダムに抽出した(表12)。

■表12: 抽出された100エッセイの内訳

スコア(レベル)	L2エッセイ(n=90)				L1エッセイ(n=10)	
	2(下位)	3(中位)	4(上位)	5(上位)	3	6
抽出したエッセイ数	20	35	30	5	5	5

(注)エッセイ内には語彙のスペルミスがしばしば存在したが、それらが本研究の結果に影響を与えることを避けるため、スペルミスのみ訂正を行った上で調査を実施した。

5.5 エッセイ評価シート

エッセイ評価シート(資料1)はいくつかの先行研究(e.g., Plakans, & Gebril, 2017)に基づいて本研究者が自作した。5つの言語的指標(一貫性・

結束性・人称代名詞・指示代名詞・つなぎ語)に関して5段階評価(1点 = Very unnatural or No – 5点 = Natural or Yes)とした。評価項目は以下の通りである。

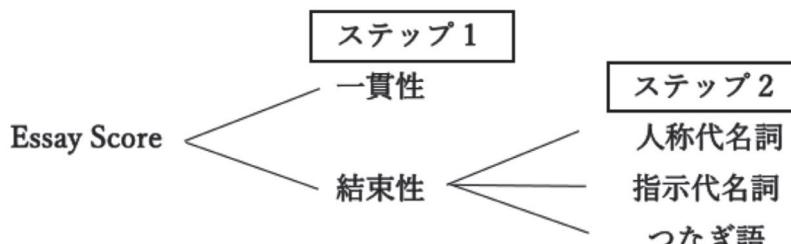
- 一貫性(Coherence): Is the logical flow of ideas natural?
- 結束性(Cohesion): Do overall cohesive devices work naturally (effectively) in this essay?
- 人称代名詞(Personal pronouns): Is the use of personal pronouns (e.g., I, you, we) natural?
- 指示代名詞(Demonstrative pronouns): Is the use of demonstrative pronouns (e.g., this, that) natural?
- つなぎ語(Connectives): Is the use of connectives (e.g., however, because, and, so) natural?

5.6 分析方法

分析には重回帰分析(Multi Regression Analysis)・順序ロジスティック回帰分析(Ordinal Logistic Regression Analysis)の2種類を採用した。重回帰分析は、1つの結果(Essay Score)に対して予測される要因が複数ある場合、その相関や影響の度合を測る標準的な分析方法である。また、順序ロジスティック回帰分析は、予測される要因が3種類以上で、かつデータが順序尺度(1~5など)に適しているため、こうした2種類の統計分析によって、より客観的な視点で結果を導き出すことにした。分析は2段階に分け、ス

テップ1ではEssay Scoreと一貫性・結束性との関連性を調査した。さらに、ステップ1で結果が得られた場合にはステップ2に進み、結束性と3つの言語的指標(人称代名詞・指示代名詞・つなぎ語)との関連性を調査した(図13)。分析には統計分析フリーソフトR version 3.5.2 (R Core team, 2019)を使用した。

本分析の前には、データの信頼性を確認するため、評価者から得られた5段階の評価点を用いてクロンバッック α 係数を求めた。その結果 $\alpha = .78$ (95% CI [.70, .85])となり、3名の評価者間信頼性(inter-rater reliability)は高いと判断した。



■図13: 調査2の分析方法

6 調査2の結果と考察

6.1 重回帰分析(Multi Regression Analysis)の結果

表13は、評価者によって評価された5つの言語的指標の記述統計である。ステップ1では、目的変数にEssay Score、説明変数に一貫性・結束性を採用した。記述統計量、各変数間の相関、散布図を確認し、VIF値より多重共線性の問題は

無かった。結果は表14の通りであり、一貫性・結束性のどちらもEssay Scoreを説明する結果となったため、ステップ2に進んだ。ステップ2は目的変数に結束性、説明変数には人称代名詞・指示代名詞・つなぎ語の3指標を採用した。いくつか競合モデルを検討した結果、得られたモデルでは指示代名詞・つなぎ語、そしてそれらの交互作用が有意であった(表15)。これらは結束性評価の65%を説明しており、予測力は高いといえる。

■表13: 評価された5つの言語的指標の記述統計

言語的指標	サンプル数	平均値	標準偏差	中央値	最小値	最大値	標準誤差
一貫性	300	2.90	1.40	3	1	5	0.08
結束性	300	3.33	1.11	3	1	5	0.07
人称代名詞	300	3.82	1.16	4	1	5	0.07
指示代名詞	300	3.48	1.27	3	1	5	0.08
つなぎ語	300	3.46	1.18	3	1	5	0.07

(注)1点=Very unnatural, 3点=Unnatural, 5点=Natural

■表14: Essay Scoreを目的変数とした、重回帰分析の結果(n=300)

言語的指標	B	95%CI	標準偏差	β	95%CI	t値	p値
一貫性	.16	[.06, .27]	.05	.22	[.08, .37]	3.13	<.001**
結束性	.22	[.09, .35]	.07	.24	[.10, .38]	3.37	<.001**

(注) ***p=0, **p<.001, *p<.01, B(偏回帰係数), β (標準偏回帰係数), 調整済み $R^2=.18$

■表15: 結束性を目的変数とした、重回帰分析の結果(n=300)

言語的指標	B	95%CI	標準偏差	β	VIF値	t値	p値
人称代名詞	.08	[.02, .17]	.04	.08	1.48	1.89	.05
指示代名詞	.38	[.36, .52]	.04	.44	1.50	10.52	<.001**
つなぎ語	.41	[.38, .53]	.04	.43	1.63	9.91	<.001**
指示代名詞・つなぎ語	.07	[.04, .17]	.03	.10	1.44	2.51	.01*

(注) ***p=0, **p<.001, *p<.01, B(偏回帰係数), β (標準偏回帰係数), 調整済み $R^2=.65$

6.2 順序ロジスティック回帰分析(Ordinal Logistic Regression Analysis)の結果

順序ロジスティック回帰分析は、6.1の重回帰分析と同様の手順で行い、使用した目的変数、説明変数も同じである。ステップ1で採用した目的変数はEssay Score、説明変数は一貫性・結束性である。記述統計量、各変数間の相関、散布図を確認し、多重共線性の問題も無かった。結果

は表16の通りであり、一貫性・結束性のどちらもEssay Scoreを説明する結果となった。よって、ステップ2に進み、目的変数に結束性、説明変数に人称代名詞・指示代名詞・つなぎ語の3指標を採用した。いくつか競合モデルを検討した結果、得られたモデルでは指示代名詞とつなぎ語の交互作用のみが有意であった(表17)。この交互作用が有意であるという結果は重回帰分析と同じであった。

6.1, 6.2の結果より、指示代名詞・つなぎ語の交互作用と結束性評価に有意な関係性が示されたため、さらに重回帰分析による交互作用効果を検討した。図14から、指示代名詞・つなぎ語と結

束性評価には正の相関があることが示され、指示代名詞・つなぎ語双方の評価が上昇するに伴って、結束性の評価が良くなることが分かった。

■表16: Essay Scoreを目的変数とした、順序ロジスティック回帰分析の結果(n=300)

言語的指標	係数	標準偏差	VIF値	z値	p値
一貫性	0.09	.00	< 2	50.85	<.001**
結束性	0.14	.00	< 2	82.11	<.001**

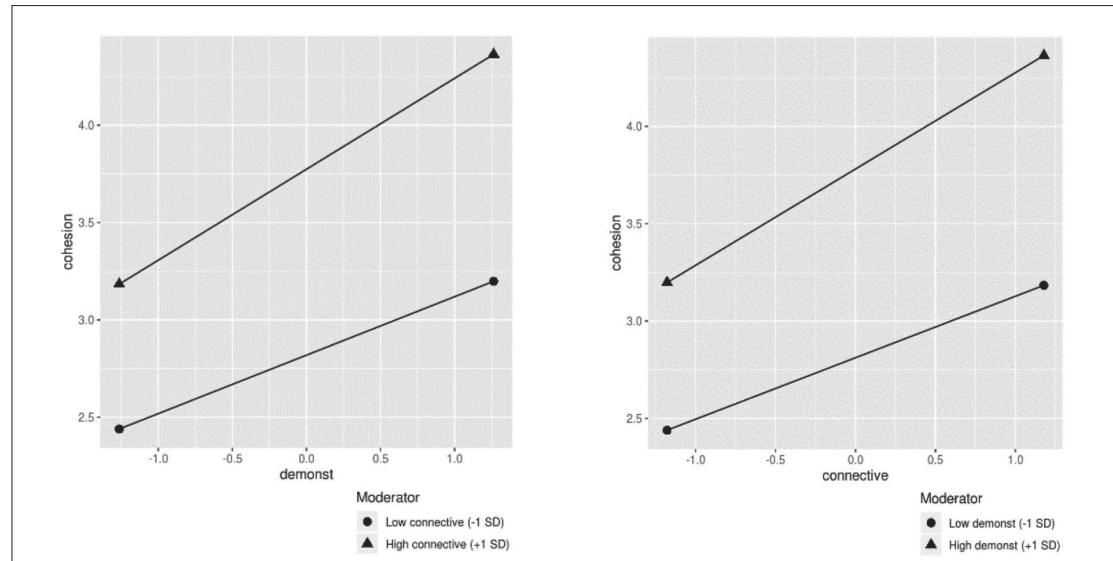
(注) ***p=0, **p<.001, *p<.01, 投入モデル clmm(essay score ~ cohesion + coherence + (1 | ID) + (1 | rater), data=dat)

■表17: 結束性を目的変数とした、順序ロジスティック回帰分析の結果(n=300)

言語的指標	係数	標準誤差	VIF値	オッズ比	z値	p値
人称代名詞	.26	.12	< 2	1.38	2.30	.02
指示代名詞	-0.13	.29	< 2	0.36	-0.47	.64
つなぎ語	.02	.30	< 2	0.45	0.09	.93
指示代名詞・つなぎ語	.42	.09	< 2	1.74	4.69	<.001**

(注) ***p=0, **p<.001, *p<.01,

投入モデル clmm(cohesion ~ personal + demonst + connective + demonst*connective + (1 | ID) + (1 | rater), data=dat)



■図14: 指示代名詞(demonst)・つなぎ語(connective)と結束性(cohesion)との関係性

6.3 アンケート・インタビュー調査の結果

学習者のエッセイの問題点を質的側面から検証するため、エッセイ評価が終了したのち、評価者にアンケート・インタビュー調査を実施した。最初に、「学習者のエッセイの問題点は何か(質

問1)」を聞いたところ、回答からは「一貫性に問題がある」という意見が多数であったため、続いて「一貫性が欠如している要因は何か(質問2)」「どこで一貫性が欠如しているか(質問3)」と一貫性に重点を置いて質問を行った。

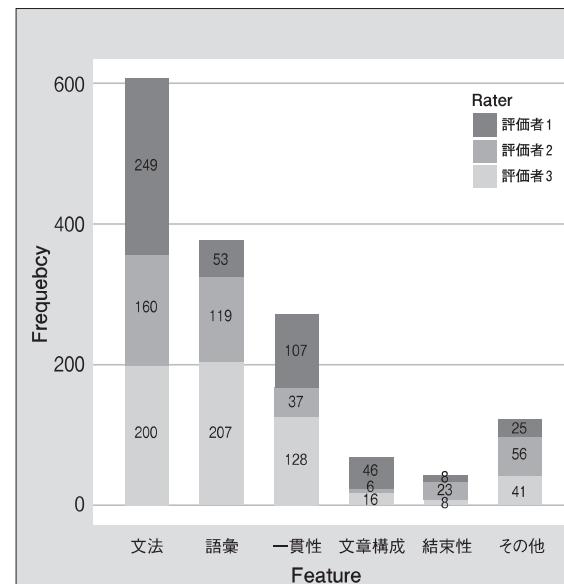
6.3.1 質問1「学習者のエッセイの問題点は何か」

質問1では「エッセイを総合的に判断した際の問題点は何か」を「一貫性」「結束性」「文法」「語彙」「文章構成」「その他」の6つの選択肢から判断するように評価者に依頼した。問題点の上位3点までを選択可能とし、得られた回答は重みづけ手法(1位⇒3倍、2位⇒2倍、3位⇒1倍と加重し、足し合わせる)を用いて算出し、積み上げグラフとした(図15)。この結果、学習者の問題点として、文法、語彙、一貫性の順に挙げられ、この結果は先行研究を支持するものであった。一方、結束性の問題点を指摘した回答は少ない。これは評価者にとって文法、語彙、一貫性のミスは判断が容易で、問題点として取り挙げやすいと考えられる。また一貫性の問題が、結束性によって起きている可能性があることを、評価者が気付いていないことも推測される。よって、この結果で、結束性は問題点が少ないと解釈することは危険であり、更なる検証が求められる。

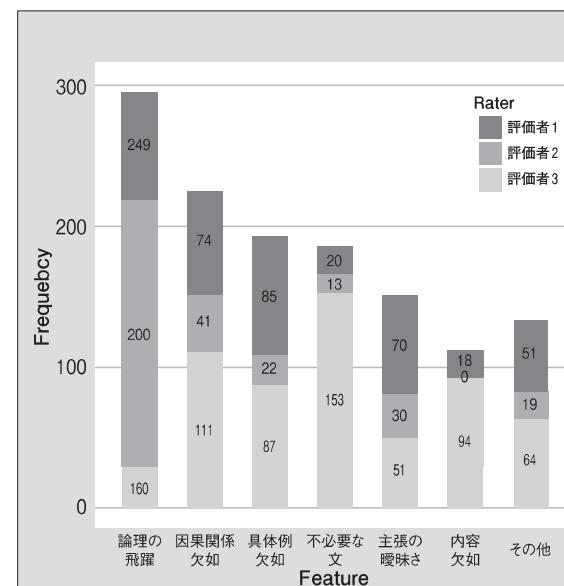
6.3.2 質問2「一貫性が欠如している要因は何か」

質問2では「論理の飛躍」「因果関係の欠如」「具体例の欠如」「不必要的文の挿入」「主張の曖昧さ」「内容の欠如」「その他」の7つの選択肢から一貫性の欠如を想起させた理由に絞って回答を依頼した。算出方法は質問1と同様である。図16の結果から、論理の飛躍が最も多く、続いて因果関係の欠如、具体例の欠如、不必要的文の挿入であった。しかし、グラフが示すように、評価者間で回答にはばらつきがあり、一貫性の欠如を感じる要因には、読み手側の背景知識量や主観的判断も影響し、個人差が生じると見えるだろう。また選択肢の各項目は明確な区別が難しく、評価者が回答を選択する際に躊躇した可能性も考えられる。

評価者によれば、「今回のエッセイ執筆者は母語(日本語)において論理的に考え、文章を組み立てる能力は十分に備わっているが、それを英語で上手く表現できずに苦戦していることがエッセイから読み取ることができ、非常にもどかしさを覚えた」と回答している。また、「母語で考えた内容をそのまま英語に当てはめようとしたり、難易度の高い語彙を用いて難しく表現しようと



■図15: 学習者のエッセイの問題点は何か？



■図16: 一貫性が欠如している要因は何か？

する傾向が伺え、むしろ平易な語彙を用いて表現力を磨いた方が良い」、さらに、「つなぎ語が文章内で上手く機能していない。つなぎ語の役割をよく理解することが大切だ」等のコメントを書いた評価者もいた。

6.3.3 質問3「一貫性が欠如しているのはどこか」

質問3では「一貫性が欠如しているのはどこか」を「エッセイ全体」「パラグラフ間」「文間(節間)」の3つの選択肢から判断するように依頼した。図17の結果より、一貫性が欠如している場所は、文と文をつなぐ文間および、節と節をつなぐ節間、つまり1文の中で起きているとの回答が多くかった。これは文や節をつなぐ作業が上手く出来ていない可能性を示唆しており、文と文、節と節のつなぎに失敗し、その結果一貫性の欠如につながっているのではないかと推測される。その一例を6.4で示す。

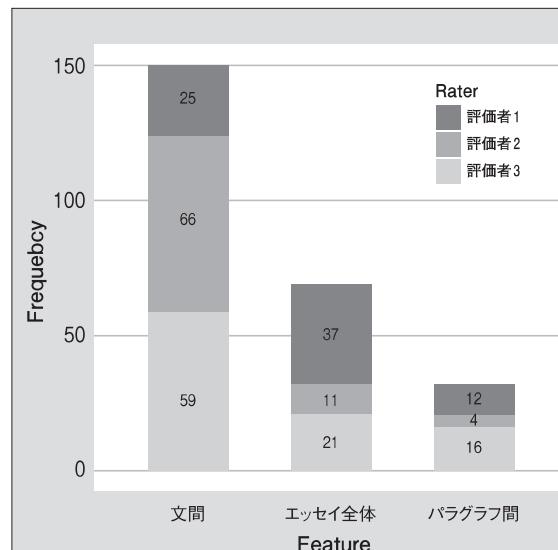


図17: 一貫性が欠如しているのはどこか?

6.4 学習者・母語話者エッセイのつなぎ表現の比較

6.3.2, 6.3.3の結果をうけ、つなぎ表現が学習者・母語話者でどのように異なるのか、具体例を以下に記す。学習者A, Bの例は、不適切なつなぎ語を使用したこと、一貫性が欠如し不自然な文章になっている。一方、母語話者Cの例は、「情報構造」(2.2参照)を意識し、指示代名詞thisを用いて文をつなぎ、首尾一貫した文章になっている例である。

6.4.1 学習者Aさんのケース

学習者Aさんが執筆したエッセイの一部抜粋である。エッセイのトピック

学習者Aさんが執筆したエッセイの一部抜粋:

In conclusion, giving enough time in their communities is necessary. There are some merits you can do. It is true that the young are so busy, then they can make a lot of time.

6.4.2 学習者Bさんのケース

次の例は学習者Bさんが執筆したエッセイの一部抜粋である。エッセイのトピック
The best way to travel is in a group led by a tour guide.
「旅行はガイド付きのパッケージツアーがよい」

Young people nowadays do not give enough time to helping their communities.

「最近の若者は地域社会に貢献する時間が十分に持てない」に対して、反対の立場で意見を述べている。

文中ではthenの前後の内容が不自然で文章に整合性がないと言える。仮にthenをhoweverに変えれば、「若者が忙しいのは事実だけれども、十分時間を作ることはできる」と文章がつながり一貫性が生じる。つなぎ語の誤用により、内容面のつじつまが合わず説明が不十分のために、論理が飛躍している例である。

に対して、賛成の立場で意見を述べている。

しかし、soの前後の内容が不自然である。節と節をsoでつなぐ場合、原因・結果の因果関係で結ぶ必要性があるが、そうしたつなぎ語の働きを十分に理解できていないと考えられる。

学習者Bさんが執筆したエッセイの一部抜粋：

Tour companies provide safety tours. They closely design a tour, therefore you won't mistake roads. Also, some foreign countries are dangerous, so you must choose your hotel, restaurant and shop when you travel by yourself.

6.4.3 母語話者Cさんのケース

次の例は母語話者Cさんが執筆したエッセイの一部抜粋である。学習者Bさんと同様のトピック「旅行はガイド付きのパッケージツアーがよい」に対して、賛成の立場で意見を述べている。

「ツアーは現地の人とコミュニケーションを取る必要が無く、それは新しい文化を理解する障壁にはなり得るけれども、しかしながら…」と続けてパッケージツアーの利点を述べようとして

いる。2文目の文頭Thisは前文全体を指しており、旧情報⇒新情報の順で情報が追加されている。

こうした文頭における指示代名詞this, that等により、過度につなぎ語を使用しなくとも、自然な流れで文をつなぐことができているため、母語話者には「情報構造」(2.2参照)への意識が備わっており、その必要性を理解していると推測される。こうした表現は、学習者にはほとんど見られなかった。

母語話者Cさんが執筆したエッセイの一部抜粋

While on the tour, most people do not need to even try to communicate outside their native language.

旧情報

新情報

This also limited their understanding and ability to engage with the new culture that they are in.

旧情報

新情報

However....

6.5 言語的指標のレベル別比較

3名の評価者によって評価された5段階の言語的指標を用いて、レベル別に平均値を算出し比較を行った(表18)。この比較にはエッセイ全体の自然さ評価(評価シートの#1の項目)を含めている(資料1)。図18の結果からも、L2は一貫性に問題があることが分かる。またL1・L2の評価点には大

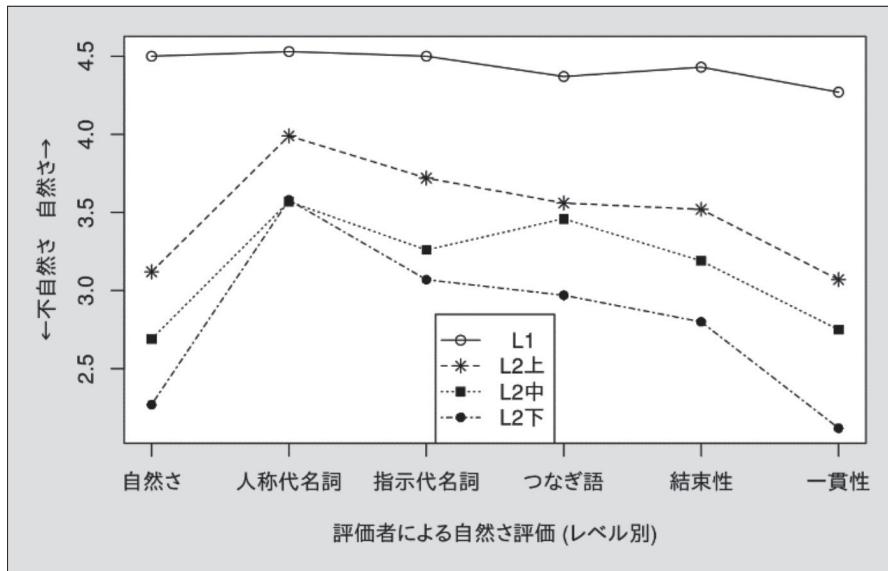
きな開きがあり、L1は全ての指標で高評価を得ている。一方、L2は下位・中位・上位と熟達度の上昇に伴い、各指標の評価が上がっている。こうした結果は、高熟達度になればなるほど、より自然な表現力が身につくことを示しており、継続的な学習の必要性が示唆された。ただし、この結果が、継続的な学習のどのような効果によるものなのかは明らかになっておらず更なる検証が必要である。

■表18: 言語的指標のレベル別比較(3名の評価点の平均値に基づいて)

レベル	全体(自然さ)	人称代名詞	指示代名詞	つなぎ語	結束性	一貫性
L1	4.50	4.53	4.50	4.37	4.43	4.27
L2上位	3.12	3.99	3.72	3.56	3.52	3.07
L2中位	2.69	3.57	3.26	3.46	3.19	2.75
L2下位	2.27	3.58	3.07	2.95	2.80	2.12

(注)1点=Very unnatural, 3点=Unnatural, 5点=Natural

本調査は、学習者がNative-likeに近づくことを目標としているわけではありません。



■図18: 6つの言語的指標の比較(レベル別)

6.6 調査2のまとめ

課題2-1は、評価者によって評価された一貫性・結束性とEssay Scoreに関連性はあるかであった。重回帰分析・順序ロジスティック回帰分析の結果より、一貫性・結束性のどちらもEssay Scoreと有意な関連性があると示された。

課題2-2は、2-1の結果が得られた場合、結束性に影響を及ぼしている具体的な指標は何か、であった。分析結果より、指示代名詞・つなぎ語の交互作用が有意であると示されたため、こうした語句を上手く組み合わせて適切に使用することが結束性に重要であると分かった。

課題2-3は、評価者へのアンケート・インタビュー調査から、学習者が抱えている一貫性・結束性の問題点は何かであった。学習者のエッセイは論理の飛躍・因果関係の欠如など一貫性に問題があり、こうした問題点は文や節をつなぐ際に生じていると回答しているため、つなぎ語が一貫性に影響を及ぼしている可能性が示唆された。

7 結論と今後の課題

本研究は、ライティングにおける文章の一貫性・結束性に焦点を当て、日本人英語学習者が抱える問題点の所在を明らかにすることであった。調査1では、日本人英語学習者112名と、英語母語話者20名によって執筆されたのべ1,020エッセイを比較検証した。その結果、結束性に関する学習者の特徴は、主語の人称代名詞、文頭・節間ににおけるつなぎ語の過剰使用であった。これは初級学習者ほどつなぎ語を多く使用する(McNamara et al., 2013)や、つなぎ語の使用頻度が低い程、熟達した書き手である(Crossley et al., 2011)と報告する先行研究を裏付ける結果となった。特に, first, second, in conclusion, for example, becauseの過剰使用に関しては、日本の高等学校や大学等でエッセイ執筆の際に使用が推奨されており、学習者の学校教育における指導も影響していると推測された。評価者に確認したところ、first, secondの使用は全く問題では無いが、少し childish, immatureな印象を受けると回答しており、実際母語話者にこれらの使用はほとんどみられなかった。これは学習者への指導が典型的なつなぎ語を教えることに限定され、つなぎ表現の多様性については指導が十分ではないことを示唆している。一方、不足に関しては、文頭つなぎ語のthen, alsoや節間のas、さらに、主語のthisであり、母語話者には容易でも、学習者には適切な使用が難しいつなぎ語があることも

確認された。また、学習者は人称代名詞 I, you, we を過剰に使用しており、特に I think が多用されていた。Academic writing における I の使用は適切ではないと従来考えられてきたが、近年では筆者のアイデンティティーを示すものとして I の使用を肯定的にとらえる研究者もおり (Hyland, 2002)、統一した見解は得られていらない。現に母語話者のエッセイ内にも人称代名詞の使用が見受けられたが、母語話者の特徴は I の使用に限定されず、多様な語彙を用いて表現ができている点である。一方、学習者は、I を繰り返し使用する傾向があり、学習者の語彙力・表現力の不足が浮き彫りになった。さらに内容面における論理の飛躍や、因果関係の欠如など、つなぎ語の誤用によって一貫性が損なわれていることも明らかとなった。

調査2は、3名の評価者によって評価された一貫性・結束性指標と Essay Score との関連性を分析し、学習者の言語的特徴がエッセイ評価にどう影響を及ぼしているのかを検証した。300 データ (100 エッセイ × 評価者 3 名分) を用いて重回帰分析・順序ロジスティック回帰分析を行った結果、一貫性・結束性のどちらも Essay Score と有意な相関が示された。さらに、結束性に影響を及ぼしている具体的な指標は指示代名詞・つなぎ語の交互作用であると示されたため、これらを上手く組み合わせて、適切に使用することがエッセイ評価にプラス (positive) の効果をもたらすことが分かった。

本研究の結果より得られた教育的示唆としては、つなぎ語の過剰使用・誤用によって一貫性が損なわれること - "Coherence can be achieved through cohesion" を理解する必要がある。よって、教育者の立場であれば、その要(かなめ)の役割である「つなぎ」が重要であることを学習者に伝え、つなぎ語の過剰使用や誤用によって起こる失敗例・成功事例を提示しながら明示的に文をつなぐ練習を行い、多様な表現力を磨かせることが学習者の助けになるだろう。また、指示代名詞等を用いた「情報構造」に意識を向けさせ、読み手を想定し「どうしたら情報が適切に相手に伝わるか」を考えさせることも必要である。熟達度が上がるほど表現が自然に近づく点を考慮すれば、継続的な指導が求めら

れるであろう。

最後に、本研究における限界点として、調査1では2つの算出方法を用いて、1エッセイあたり・200語あたりで検証を試みたが、これには議論の余地がある。例えば、first, second などエッセイ特有の表現を200語あたりで算出する必要性は感じられず、むしろ1エッセイあたりでどの程度使用頻度率があるかを算出することに意味があるだろう。一方、and, but などエッセイに限らず最頻出される語は200語あたりで調査するほうが妥当性があると言える。どのような検証方法が適切であるかはさらなる可能性を模索する必要がある。また、調査2では文法・語彙の項目を研究対象に含めなかった。しかし、それらを含めた包括的な分析から、より客観的な結果を導く必要性があったと感じている。今後の研究ではこうした限界点を考慮し、言語的指標を追加してエッセイ評価との関連性をより多角的な視点で捉え、ライティングにおける一貫性・結束性研究に新たな示唆を与えることが望まれる。本研究が日々努力を重ねる日本人英語学習者のライティング力向上の一助になれば幸いである。

謝 辞

本研究を実施・発表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さま、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、助言者である寺内一先生には、有益なご助言・ご指導をいただきましたこと、深く感謝申し上げます。そして名古屋大学の杉浦正利先生には、本研究の立案から実施、報告書執筆に至るまで、お忙しい中たくさんの貴重なご教示と温かな励ましの言葉をいただきました。この場をお借りして心より感謝申し上げます。また大学院で共に学んだ山田敦子さんにも多くのサポートをいただきましたこと、深く御礼申し上げます。最後に、研究協力に快諾して下さった母語話者の Kelly さん, Don さん, Dimy さんと本研究に携わって下さった全ての皆様に厚く御礼申し上げます。

参考文献 (*は引用文献)

- * Bloor, T., & Bloor, M. (2013). *The functional analysis of English: A Hallidayan approach. Third Edition*. Routledge.
- * Burstein, J., & Wolska, M. (2003, April). Toward evaluation of writing style: finding overly repetitive word use in student essays. In *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 1* (pp. 35-42). Association for Computational Linguistics.
- * Carrell, P. L. (1982). Cohesion is not coherence. *TESOL quarterly*, 16 (4), 479-488.
<https://doi.org/10.2307/3586466>
- * Collins, J. L. (1998). *Strategies for struggling writers*. Guilford Press.
- * Crossley, S., & McNamara, D. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 32 (32).
- * Crossley, S. A., Weston, J. L., Sullivan, S. T. M., & McNamara, D. S. (2011). The Development of Writing Proficiency as a Function of Grade Level: A Linguistic Analysis. *Written Communication*, 28 (3), 282-311.
- * Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). 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.
- * Daud, N. S. M., Daud, N. M., & Kassim, N. L. A. (2016). Second language writing anxiety: Cause or effect? *Malaysian journal of ELT research*, 1 (1), 19.
- * Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of second language writing*, 4 (2), 139-155.
- * Ersanlı, C. Y. (2015). Insights from a Learner Corpus as Opposed to a Native Corpus about Cohesive Devices in an Academic Writing Context. *Universal Journal of Educational Research*, 3 (12), 1049-1053.
- * Farghal, M. (1992). Naturalness and the notion of cohesion in EFL writing classes. *International Review of Applied Linguistics in Language Teaching*, 30 (1), 45- 50.
- * Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing. An Applied Linguistic Perspective*. Routledge.
<https://doi.org/10.4324/9781315835853>
- * 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, & computers*, 36 (2), 193-202.
- * Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Routledge.
- * 平成29年度英語教育改善のための英語力調査 事業報告. 平成29年度 英語力調査結果(高校3年生).
https://www.mext.go.jp/a_menu/kokusai/gaikokugo/_icsFiles/afieldfile/2018/04/06/1403470_03_1.pdf
- * Hopp, H., Bail, J., & Jackson, C. N. (2020). Frequency at the syntax-discourse interface: A bidirectional study on fronting options in L1/L2 German and L1/L2 English. *Second Language Research*, 36 (1), 65-96.
- * Hyland, K. (2002). Options of identity in academic writing. *ELT Journal*, 56 (4), 351-358.
<https://doi.org/10.1093/elt/56.4.351>
- * Kojima, M. (2019). Metadiscourse in Japanese EFL Learners' Argumentative Essays: Applying the Interpersonal Model. *The Asian EFL Journal*.
- * Lee, I. (2002). Teaching coherence to ESL students: A classroom inquiry. *Journal of second language writing*, 11 (2), 135-159.
- * Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, 33 (4), 623-636. <https://doi.org/10.1016/j.system.2005.02.002>
- * Lorenz, G. (1998). Overstatement in advanced learners' writing: stylistic aspects of adjective intensification. In. S. Granger (Ed.) *Learner English on computer* (pp. 53-66).
- * Luo, D., Gu, W., Luo, R., & Wang, L. (2016). Investigation of the effects of automatic scoring technology on human raters' performances in L2 speech proficiency assessment. *Chinese Spoken Language Processing*, 1-5.
- * Machado, C., & Tuero, S. (2013). Do Students Use More Connectors When They Write Longer Texts? A Study on the Use of Connectors in Paragraphs and Essays Written by EFL College Majors. *ACTUALIZACIÓN EN LA ENSEÑANZA DEL INGLÉS*, 103.
- * Master, P. (2002). Information structure and English article pedagogy. *System*, 30(3), 331-348.
- * McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45,499-515.
- * McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.
- * Narita M. (2018). How We Can Foster Japanese EFL Learners' Writing Proficiency. *Learner Corpus Studies in Asia and the World*, 3:43-60. http://www.lib.kobe-u.ac.jp/handle_kernel/81010117
- * Natsukari, S. (2012). Use of I in essays by Japanese EFL learners. *JALT Journal*, 34 (1), 61-78.
- * Oi, K. (2002). Writing Instruction Aiming to Foster Student' Thinking Faculty : From the Point of View of EFL II. *Humanities and Social Sciences, the 50th Anniversary Edition*.
- * Oi, K. (2008). Fostering Junior High School Students' Ability to Think through Teaching Writing in English. 千葉大学教育学部研究紀要 第56巻175~184頁 <https://core.ac.uk/reader/96989795>
- * Pinker, S. (2015). *The Sense of Style: The thinking person's guide to writing in the 21st century*. Penguin Books.
- * 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. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

参考文献(*は引用文献)

- * Todd, R. W., Thienpermpool, P., & Keyuravong, S. (2004). Measuring the coherence of writing using topic-based analysis. *Assessing Writing*, 9 (2), 85-104.
- * Todd, R. W., Khongput, S., & Darasawang, P. (2007). Coherence, cohesion and comments on students' academic essays. *Assessing Writing*, 12(1), 10-25.
- * Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech, Language, and Hearing Research*, 38 (6), 1349-1355.
- * Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education*, 23 (1), 31-48. <https://doi.org/10.1016/j.linged.2011.09.004>
- * Zarifi, A., & Mukundan, J. (2014). Creativity and unnaturalness in the use of phrasal verbs in ESL learner language. 3L: *Language, Linguistics, Literature*®, 20 (3).

資料1：エッセイ評価シート

A study on learners' persuasive essays						
<p>The following questions will ask you (native speakers) about persuasive essays written by Japanese learners of English. First, please read their essays. Then, check or write the index that most closely indicates the extent of English fluency and naturalness which you perceive in the essay:</p>						
Appropriate response to each question below:						
#	Question	1= Very unnatural (strongly disagree) 3= Unnatural (disagree) 5= Natural (agree)				
1	Is this essay's use of English natural?	1	2	3	4	5
2	The biggest issues are... A: Vocabulary B: Connectives C: Logical flow of idea D: Grammar E: Organization F: Other	1st: 2nd: 3rd: (if you have)				
3	Is the use of personal pronouns (Ex: I, you, we) natural?	1	2	3	4	5
4	As a whole, what is major reason? (No need to answer this, if 1 in #3.) A: Overuse B: Underuse C: Misuse D: Misplace E: Other	(Choose one)				
5	Is the use of demonstrative pronouns (Ex: this, that) natural?	1	2	3	4	5
6	As a whole, what is major reason? (No need to answer this, if 1 in #4.) A: Overuse B: Underuse C: Misuse D: Misplace E: Other	(Choose one)				
7	Is the use of connectives (Ex: however, because, and, so, firstly, therefore, in conclusion) natural?	1	2	3	4	5
8	As a whole, what is major reason? (No need to answer this, if 1 in #5.) A: Overuse B: Underuse C: Misuse D: Misplace E: Other	(Choose one)				
9	Do overall cohesive devices work effectively in this essay?	1	2	3	4	5
10	Is the logical flow of ideas natural?	1	2	3	4	5
		Yes	No			
		1	2	3	4	5
		1st: 2nd: 3rd:				
		(Choose one)				
		1	2	3	4	5
		Yes	No			
THANK YOU for your cooperation!						
本研究では # 10 は使用していません。						

A 研究部門・報告Ⅳ・英語能力テストに関する研究

多様な言葉の働きを引き出す タスクタイプの組み合わせ — 言語機能分析による提案 —

研究者:茨城県／筑波大学大学院 在籍 前田 啓貴

《研究助言者:和田 稔》

概要

多様な発話を学習者から引き出すことを目的に指導や評価の場面でコミュニケーションタスクが使用されるようになってきている。本研究では、コミュニケーションタスクで代表的なジグソー、情報ギャップ、問題解決、意思決定、意見交換、ナレーションの6種類のタスクタイプの特徴を、タスク中に引き出される言語機能の観点から調査した。その結果、与えられたものを描写するタイプ(ナレーション、ジグソー、情報ギャップ)と意見を考えて伝えるタイプ(問題解決、意思決定、意見交換)のタスクに分類することができた。さらに、同じ「描写」型のタスクと「意見」型のタスクの中でも、引き出される言語機能の数がタスクタイプによって異なり、最低で1種類、最高で9種類の言語機能を1つのタスクで引き出すことが可能であった。本研究の結果から、タスクを使用する際に、どのようにタスクを選び、組み合わせるかについて、言語機能を考慮した方法を提示した。

1 はじめに

近年の英語教育では、コミュニケーションツールとして英語を使うことが重要視され、ペアやグループによるスピーキング活動が指導や評価の場面で導入されるようになってきている。例えば、授業では、一方向の発話であるスピーチやプレゼンテーションだけではなく、双方向のやりと

りであるディスカッションや情報ギャップを取り入れた活動が散見される。また、評価においても、パフォーマンステストの一環として、スピーキングの評価が実施されるようになってきている。ペアやグループでのスピーキングでは、会話を維持または発展させていくことが求められるため、インタビューなどのモノローグ型のスピーキングとは異なり、会話をするために多くの能力が必要になる。

このような場面では、言語活動の一種であるタスクを中心に活動やテストが構成され、学習者は指定された言語形式の意味理解やその产出を目的にするのではなく、それぞれの課題に設定されたゴールを達成することが求められる。そして、タスクのゴールを達成するまでのプロセスで、学習者はコミュニケーションの目的や場面に合わせて言葉を使用することになる。タスクに基づく言語指導(Task-based language teaching)に関する研究においても、タスクを用いることで様々な学習効果が期待できることがわかっている(e.g., Ellis, 2003)。特に、言語の意味に焦点を当てた有意味なやりとりを行う機会を与える手段の一つとしてタスクは有効であり、今後さまざまな場面で活用されると考えられる。

しかし、教科書等の単元末に掲載されている言語活動は、学習者の注意を形式に向けるよう、あくまで特定の形式を产出させることを目的としている活動が多いことが指摘されている(藤田, 2020)。教室で実際にタスクに基づく指導や評価を導入する際には、単元末にある活動を工夫

することや、新しくタスクを作成、または他の教材から選ぶ必要があり、これらはタスクを導入することを難しくする一因になっていると考えられる。

また、コミュニケーションタスクは教室規模の評価から大規模テストまで、幅広く利用されている。大規模なテストでは、学習者が目標言語について何を知っているかではなく、目標言語を使って何ができるかに焦点を当てて評価することが主流になっており(Wiggins, 1998)、学習者のパフォーマンスを引き出すためのタスクは重要な役割を果たしている。大規模テストは、評価の目的がはっきりしているため、使用されるタスクが頻繁に変更されることはほとんどない。

一方で、教室で行われる評価では、これまでに学習してきた内容とテスト内容の一貫性について、評価を行うたびに確認するだけでなく、テストにおけるパフォーマンスと教室の外で学習者が実際にできることのつながりを考慮する必要がある。大規模なテストとは少し異なり、このように、教室で行われる評価にタスクを導入する際には、どのようなタスクをテストに含めるかを決めるという難しさがある。

そこで本研究では、学習内容、評価、そして実際に学習者ができることの一貫性を確保するための視点として言葉の働きに着目し、コミュニケーションタスクとして代表的な6種類のタスクタイプの特徴を、タスク中に使用される言葉の働きの観点から明らかにする。本研究の結果から、指導と評価におけるタスクの利用について示唆を与えることを目指す。

2

先行研究

2.1 タスクとタスクタイプ

タスクとは、目標を達成するために学習者が言語を使用する必要がある活動のことであり、与えられた課題に対して、正確または適切な内容が伝えられたかという、言語の意味に焦点が当てられた活動のことである(Ellis, 2003)。タスクは、あ

らかじめ決められた特定の言語形式(e.g., 文法項目)を用いて課題を遂行させることを目的とした焦点化タスク(focused task)と、学習者が既にもっている言語または非言語的なリソースを用いてコミュニケーションを行うことを促す非焦点化タスク(unfocused task)に分けることができるが、原則として以下の4つの条件を満たしていなければならない(Ellis, 2009)。

1. 言語の意味に主な焦点が当てられること。
2. 学習者間に、意見の相違やもっている情報の違いなどのギャップが存在すること。
3. 学習者は、タスクを行うために自身がもつ言語または非言語的リソースを利用すること。
4. 課題に関連した明確なゴールがあること。

まず、1つ目として、タスク中のやりとりは、例示されているダイアログの模倣や与えられた文から単語を入れ替えるような言語操作が目的ではなく、学習者がもつ内容や考えを伝えたり、理解したりすることが目的でなければならない。2つ目に、意味に焦点を当てたやりとりを行うためには、学習者間で少しでも異なる意見をもっていたり、課題を解決するために必要な情報が分散した状態で学習者に与えられているなど、何かしらのギャップが存在することが必要である。そして、3つ目は、学習者がタスク中に使用する言語が、事前に教師から形式等の制限をされないことを意味している。これは、焦点化タスク(focused task)においても同様に求められ、教師は学習者に特定の言語形式等を使用することを期待しているとしても、それを明示的に使用するように指示することはない。つまり、学習者は、タスク中に自分が既に持っている言語または非言語的リソースを用いて、ゴール達成に必要なコミュニケーション上の役割を果たすことが求められる。最後に、タスクは、文法的に正しく英語を産出できるかどうかといった言語そのものに関するゴールではなく、それぞれのタスクに関連するゴールを達成するために言葉を使う活動でなければならない。

コミュニケーションタスクはこれまで、様々なものが考案されてきたが、それらは大きく分けて(a) ジグソー、(b) 情報ギャップ、(c) 問題解

決, (d) 意思決定, (e) 意見交換, (f) ナレーションの6種類に分けることができるとされる(松村, 2017a)。松村は、それぞれのタスクを、「タスク中に学習者が期待される行為」という観点からタスクタイプとして分類しており、タスクを準備す

る際には、このタスクタイプを考慮することが重要であると主張している(松村, 2017b)。これら6種類のタスクの定義と代表的なタスクの例、そして活動内容をまとめたものが表1である。

■表1: タスクタイプとその定義(松村, 2017aを改変)

タイプ	定義	タスク例	活動内容
ジグソー (jigsaw)	学習者がそれぞれ持っている情報を合成することによって全体像を完成させる。	物語復元	物語を構成する一連の絵のフレームが学習者に均等に配分され、学習者は互いに情報を与え合って元の物語を再構築する。
情報ギャップ (information-gap)	相補的な情報を持つ学習者が互いに相手から情報を得て課題を解決する。	相違特定	相似しているが同じではない2枚の絵の間の違いを、学習者どうし絵を見せあうことなく特定する。
問題解決 (problem solving)	問題状況と解決のための鍵を与えられ、正答に到達する。	論理パズル	論理的分析や推論によって解決されるべき問題の答えを、学習者が協力して考える。
意思決定 (decision-making)	多様な可能性の中からの選択や、候補に対する順位づけ、解決策の提案などをする。	選択・順位づけ	学習者が各自の意見を持ち寄り、与えられた条件下で最善の選択や順位を導き出す。
意見交換 (opinion-exchange)	トピックについて各自の意見を述べ合い、討論する。	時事批評	今日の重要な社会的問題に対し、学習者がそれぞれ自分の意見を述べて議論する。
ナレーション (narration)	ストーリーを見ながら、もしくは見た後でその内容を説明する。	物語再話	一連の絵で表現されたストーリーの内容を語る。

上述したように、それぞれのタスクタイプについて、先行研究ではさまざまな定義が存在するが、とくに問題解決タイプについては、研究者間でもどのようなタスクを含めるのかが大きく異なっている。松村(2017b)では、事前に設定された正答に到達することが必要であるとし、問題状況の解決策を考えるような課題は「問題解決的」ではあるものの、問題解決タイプには含まれないとしている。一方で、Nation(2013)や紺渡(2015)、柳川(2017)では、問題解決活動または問題解決タイプの典型的な例として、解決策を考案するタスクを紹介している。例えば、Nation(2013)では、環

境汚染問題に対する解決策を提案する活動を例にあげている。本研究では、用語に対する活動内容のわかりやすさを考慮し、紺渡(2015)が提案している与えられた問題に対して理由を示して解決策を提案する問題解決型の活動も問題解決タイプに含めるとする。

指導や評価においてタスクの使用が増えてきているのに伴い、タスクタイプがタスク中の学習者の発話に与える影響について調査する研究が行われるようになってきている。例えば、タスクのゴール達成までに必要なターン数がタスクによって異なり、会話が単純なタスクと複雑なタス

クが存在することがあげられる(Pallotti, 2017)。また、テスティングの分野においては、テストタスク中に使用される言語機能を調査している研究も存在する(e.g., He & Dai, 2006)。

しかし、タスクタイプに焦点を当てた研究は、Robinson(2005)などに代表される、タスクの認知的複雑さに基づくタスクの選択と配列に関する研究に比べて少なく、タスクタイプの影響について未解明な部分が多いのが現状である。

松村(2017b)が主張するように、タスクを使用する際には、タスク中の認知プロセスに基づいて使用するタスクを判断することは難しいため、タスク中に学習者が求められる行為に着目することが重要であると考えられる。

2.2 言葉の働き

言葉は、コミュニケーションを成り立たせるために必要な、何かしらの機能を果たすために用いられる。このようなコミュニケーションの中で言葉がもつ機能は、「言葉の働き」や「言語機能」と呼ばれる。これまで、コミュニケーションの中

で適切な機能をもつ発話を産出する能力は、社会言語学的能力や語用論的能力に分類されてきた(Green, 2012)。例えば、ヨーロッパ言語共通参考枠(CEFR; Council of Europe, 2001)では、特定の目的のコミュニケーションにおいて必要な機能を果たすために言語リソースを使用する能力は、語用論的能力の一部として位置付けられている。

CEFRは、行動志向(action-oriented)の言語観をベースに開発されており、学習者の熟達度は「言葉を使って何ができるか(CAN DO)」というCAN-DOディスクリプタ(能力記述文)で記述されている。日本の英語教育においても、CEFRの取り組みを参考に、学習目標を「～ができる」の形式であるCAN-DOリストとして設定することが徐々に実行されはじめしており(小林, 2019), CAN-DOリストの作成にあたっては、表2に示すような、学習指導要領に記載されている言葉の働きを参照することが推奨されている(文部科学省, 2013)。今後、指導と評価の一体化や、異なる校種間で一貫した指導を行うために、CAN-DOリストのさらなる活用が期待されている。

■表2: 言葉の働きの例

言葉の働き	例		
コミュニケーションを円滑にする	・言い換える ・聞き直す	・話題を発展させる ・話題を変える	・相槌をうつ ・繰り返す
気持ちを伝える	・謝る ・褒める	・共感する ・望む	・驚く ・心配する
事実・情報を伝える	・説明する ・描写する	・要約する ・報告する	・理由を述べる ・訂正する
考えや意図を伝える	・提案する ・賛成する	・承諾する ・主張する	・仮定する ・断る
相手の行動を促す	・質問する ・依頼する	・誘う ・助言する	・命令する ・説得する

言葉の働きは、学習者がタスク中に言葉を使って何をしているか、または何をすることが求められるのかを調査するための手段としても用いられる。例えば、Nakatsuhara(2014)では、Test of English for Academic Purposes (TEAP) のスピーキングセクションを開発する段階で、O'sullivan, Weir, & Saville(2002)が提案した

言語機能のチェックリストを用いて、スピーキングタスク中に引き出されると事前に想定した言語機能を、タスク中に学習者が本当に使用しているかを確認している。TEAPのスピーキングセクションにおいて、受験者は、インタビュー、ロールプレイ、短いモノologueの3種類のタスクを、4つのパートの中で行う必要がある。結果として、

インタビュータスクでは、「意見を主張する」や「意見を正当化する」などの言語機能が用いられ、ロールプレイでは「挨拶をする」、「情報を尋ねる」といった言語機能が用いられるなど、3つのタスクの中で多様な言語機能を引き出すことができるこことを明らかにした。

Koizumi & Hirai (2012) では、Story Retelling Speaking Test (SRST), Pearson Education の Versant, ALC の Standard Speaking Test (SST) という3つのテスト間で、タスク中に引き出される学習者の発話がもつ機能を比較している。結果として、SST, SRST, Versant の順でより多様な言語機能を引き出すことができていることを明らかにした。また、重要な点として、Koizumi & Hirai は、より多くの言語機能を引き出せるタスクが良いとは限らず、タスクを使用する目的や、タスク使用者が意図した言語機能を受験者が実際に実行していることが重要だとしている。

He & Dai (2006) が主張するように、コミュニケーションタスクは幅広い機能を持つ発話を引き出すことが想定されているが、タスク中に学習者が本当に産出しているとは限らない。そのため、コミュニケーションタスク中に学習者がどのような言語機能をもつ発話を実際に産出しているかを明らかにすることは、学習と評価におけるタスクの使用に対して多くの示唆を与えることができると考えられる。

3 本研究

3.1 本研究の概要と目的

これまでに述べたように、コミュニケーションを重視する指導や評価を行う場面において、タスクを活用することが期待される。また、指導や評価の一体化という観点からは、CAN-DO リストの活用を理由に、今後は指導や評価における言葉

の働きの重要性が増すと考えられる。さらに、タスクと言語機能は、どちらも個別具体的な語彙や文法項目の使用が目的ではなく、学習者が言葉を使ってどのような行為ができるのかに着目しているという点で親和性が高いと考えられる。しかしながら、タスク中に用いられる発話を言葉の働きの観点から明らかにした研究は少なく、タスクと言葉の働きの関係は十分に明らかではない。

そこで本研究では、コミュニケーションタスクに代表される6種類のタスクタイプ(i.e., ジグソーや情報ギャップ、問題解決、意思決定、意見交換、ナレーション)を、タスク達成までに用いられる言葉の働きの観点から調査することを目的とした。本研究の結果から、学習や評価に使用するコミュニケーションタスクを、言葉の機能を考慮して選択できるようになり、タスクを選ぶための1つの基準を示すことができると思われる。また、複数のコミュニケーションタスクを使用する際には、言葉の働きの観点から複数組み合わせてタスクを実施することができるようになり、多様な機能をもつ発話を学習者から効率的に引き出すことができるタスクの組み合わせを提案できる。リサーチクエスチョン (RQ) は以下のとおりである。

RQ1

ジグソー、情報ギャップ、問題解決、意思決定、意見交換、ナレーションの6種類のタスクで引き出される言語機能はどのように異なるか。

3.2 方法

3.2.1 協力者

日本国内にある高校または大学に通う英語学習者24名(12ペア)が実験に参加した。協力者の実験参加時の実用英語技能検定の取得級は表3の通りである。

■表3: 協力者の英検取得級

級	1級	準1級	2級	準2級	3級	4級	未記入(なし)
n	1	1	7	4	1	1	9

3.2.2 タスク

TBLTに関する先行研究で使用または言及されているタスクから選定した。使用されている英語の難易度等の理由からそのまま使用することができないと考えられたため、タスクの細部については筆者がタスクを改変した。各タスクの概要と引用元は表4の通り。すべてのタスクは2人1組で行われた。

ジグソータイプには、物語復元タスクを採用した。参加者は4コマ漫画のうちランダムに2コマずつ与えられ、相手に自分の絵を見せることなく、もとの4コマ漫画のストーリーを完成させるように指示された。情報ギャップタイプには、Amagai(2018)で使用された間違い探しタスクを使用した。このタスクでは、参加者は、9種類いる動物のうち2種類の位置が異なっている動物園の地図が渡され、どの動物の位置がペアで異なるかを見つけるように指示が出された。問題解決タイプでは、紺渡(2015)で使用された形式を用いて、解決する問題

を変えて使用した。具体的には、歩きスマホをする人を減らすための解決策を1人1つ以上考えた上で、ペアで話し合い、最も良い解決策を提示することをタスクのゴールとした。意思決定タイプでは、ランキング作成タスクを使用した。National Aeronautics and Space Administration(n.d.)によって作成されたタスクであり、月に不時着した宇宙飛行士が、基地まで歩いて移動する際に持っていく荷物を決めるために15種類のアイテムの中で順位をつけるタスクである。本研究では、時間の都合上、すべてのアイテムに順位付けをするのではなく、上位5つをペアで話し合って順位を決めるタスクとした。ディスカッションタスクでは、「未来にも学校は存在するか」というお題について、お互いに意見を出し合い、相手の意見に賛成または反対かを伝え、その理由を述べ会うことを目的とした。最後に、ナレーションタイプは、4コマ漫画が与えられ、30秒の準備時間の後、その内容についてペアの人に伝えるというタスクを行なった。

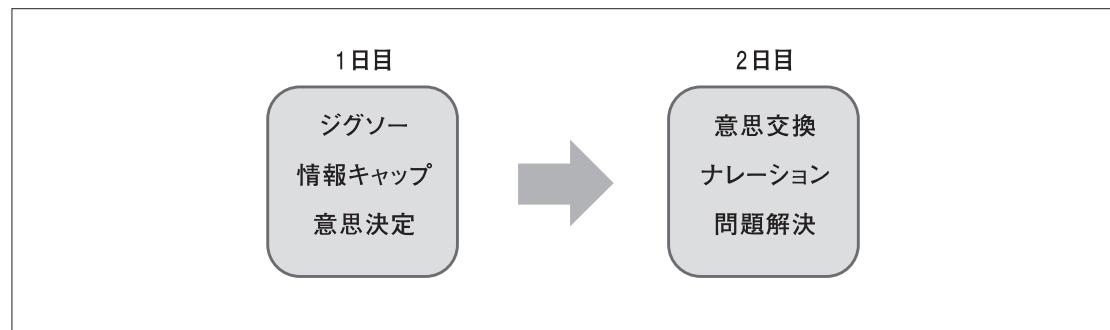
■表4: 本研究における各タスクタイプで使用したタスク及びその概要

タスクタイプ	タスク	出典や参考	概要
ジグソー	物語復元	筆者作成	4コマ漫画のうち2コマずつをペアに配布し、相手の絵を見ることなく4コマ漫画を復元する。
情報ギャップ	間違い探し	Amagai (2018)を改変	全9種類の動物のうち、2種類の動物の位置が異なる地図をペアに配布し、相手の地図を見ることなく位置が異なる動物を探す。
問題解決	解決策提示	紺渡(2015)を改変	歩きスマホをする人を減らすための解決策を、理由を含めて考える。
意思決定	ランキング作成	National Aeronautics and Space Administration (n.d.)を改変	月に不時着した宇宙飛行士が、基地まで歩いて移動する際に持っていく荷物を、15個の中から上位5つを決める。
意見交換	ディスカッション	筆者作成	「未来にも学校は存在するか」というお題について、お互いに意見を出し合い、相手の意見に賛成または反対かを伝え、その理由を述べる。
ナレーション	4コマ漫画描写	英検準1級スピーキングテストを再現	4コマ漫画が与えられ、30秒の準備時間の後、その内容を描写し相手に伝える。

3.2.3 手順

協力者は熟達度が近い他の協力者と2人1組のペアになり、図1にあるように、6種類のタスクを

2日に分けて順番に取り組んだ。タスク間でペアの変更はせず、同じペアで全てのタスクに取り組んだ。それぞれのタスクは15分以内に終了した。



■図1：実験の手順

3.2.4 分析

分析は、発話の機能を詳細に分類することができる言語機能分析を行った。録音されたタスク中の会話を書き起こし、調査者と英語教育学を専攻する大学院生1名によって全体の3割に対してコード付けを行った。問題点がある場合には協議を行って解消した。その後、協議の結果に基づき、調査者1人が残りのデータにコード付けを行った。コード付けは、Observation checklist (O'sullivan, Weir, & Saville, 2002)に基づき、各ターンに対してコード付けを行なった(各言語機能の詳細は資料1を参照)。Observation checklistは、一方向的な情報提供の機能をもつinformational function、双方向的なやりとりが必要とするinteractional function、やりとりを管理するmanaging functionの大きく分けて3種類の言語機能から構成されるチェックリストであり、それぞれのfunctionは合計で31種類の言語機能に細分化して分類することができる。1つのターンが複数の言語機能をもっている場合には、複数のコードをつけた。最終的に、全ての発話は付与された言語機能ごとに分類、集計され、言語機能ごとに、各タスク中に1人あたりが使用した回数を算出した。タスクのゴール達成のためにタスクが引き出した言語機能とみなす閾値は、Nakatsuhara(2014)に基づき1人あたり0.7回以上の機能とした。また、本研究では、タスク中にどちらか片方の参加者が使用する可能性が高い言語機能も調査するため、1人あたり0.35以上、

すなわちペアで0.7回以上の出現頻度が確認された言語機能についても確認した。

4 結果と考察

まず、6種類のタスク全体において使用された言語機能を比較し、その後、それぞれのタスクの特徴について議論する。

表5に各タスクの言語機能分析の結果を示す。まず、全てのタスクに共通して使用された言語機能は存在しなかった。タスク中にはほぼ全員が使用することが期待される0.7回以上の出現率だった言語機能は、最も少なかったナレーションタスクで1つ(i.e., 順序的に描写する)、最も多かった問題解決タスクで5つ(i.e., 意見を述べる、理由を述べる、提案する、理解を示す、決定する)であった。タスク中に少なくともペアのうちどちらかが使用することが期待される0.35回以上の出現率だった言語機能を含めると、問題解決タイプが最も多くの言語機能をタスク中に引き出すことが可能であり、合計で9つの言語機能(i.e., 意見を述べる、意見の明確化、理由を述べる、提案する、賛成する、意見を求める、理解を示す、会話を発展させる、決定する)を引き出すことがわかった。また、タスクタイプ間で共通の言語機能に着目すると、伝える内容が決まっているナレーション、ジグソー、情報キャップの3つが「描写する」や「順

序的に描写する」といった点で共通している「描写」型のタスクであり、伝える内容も考える必要がある問題解決、意思決定、情報交換の3つが「意見を述べる」で共通している「意見」型のタスクであることがわかった。また、「意見」型のタスクの方が「描写」型のタスクより多くの言語機能を引き

出すことがわかった。しかし、全体の傾向としては、各タスクが引き出す言語機能は非常に限られていた。次からのセクションでは、それぞれのタスクの特徴を、使用された言語機能の少ないタスクから順番に考察していく。

■表5: タスク中における各言語機能の出現頻度の割合と1人あたりの出現回数の平均

言語機能	タスクタイプ											
	ジグソー		情報ギャップ		問題解決		意思決定		意見交換		ナレーション	
	割合(%)	平均	割合(%)	平均	割合(%)	平均	割合(%)	平均	割合(%)	平均	割合(%)	平均
個人的な情報(現在)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
個人的な情報(過去)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
個人的な情報(未来)	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
意見を述べる	0.00	0.00	0.33	0.08	20.97	1.63	26.50	2.58	25.00	1.17	0.00	0.00
意見の明確化	2.35	0.08	0.66	0.00	6.45	0.50	2.99	0.29	8.93	0.42	0.00	0.00
理由を述べる	0.00	0.00	0.00	0.0	9.14	0.71	5.13	0.50	21.43	1.00	0.00	0.00
比較する	1.18	0.04	0.98	0.13	0.00	0.00	0.00	0.00	0.89	0.04	0.00	0.00
推測する	3.58	0.13	2.30	0.29	0.00	0.00	1.28	0.13	1.79	0.08	0.00	0.00
段階的に述べる	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.04	0.00	0.00	0.00	0.00
順序的に描写する	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.0	1.00
描写する	52.94	1.88	50.82	6.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
要約する	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	12.50	0.58	0.00	0.00
提案(主張)する	11.76	2.02	0.98	0.13	14.52	1.13	2.99	0.29	4.46	0.21	0.00	0.00
好みを伝える	0.00	0.00	0.00	0.00	1.08	0.08	0.00	0.00	0.00	0.00	0.00	0.00
賛成する	1.18	0.04	5.90	0.75	5.91	0.46	14.53	1.42	9.82	0.46	0.00	0.00
反対する	0.00	0.00	1.97	0.25	0.54	0.04	1.71	0.17	0.00	0.00	0.00	0.00
修飾する	0.00	0.00	0.33	0.04	0.00	0.00	0.00	0.00	0.89	0.04	0.00	0.00
意見を求める	0.00	0.00	0.00	0.00	6.45	0.50	13.68	1.33	4.46	0.21	0.00	0.00
説得する	0.00	0.00	0.00	0.00	0.00	0.00	2.56	0.25	0.00	0.00	0.00	0.00
情報を求める	4.71	0.17	14.75	1.88	1.61	0.13	1.28	0.13	0.00	0.00	0.00	0.00
会話の修復する	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
理解を確認する	3.53	0.13	3.61	0.46	0.00	0.00	1.28	0.13	1.79	0.08	0.00	0.00
理解を示す	4.71	0.17	5.25	0.67	9.68	0.75	6.41	0.63	2.68	0.13	0.00	0.00
話す土台を確立する	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.04	0.00	0.00	0.00	0.00
明確化を求める	2.35	0.08	2.30	0.29	1.61	0.13	0.00	0.00	0.00	0.00	0.00	0.00
他者を訂正する	0.00	0.00	0.33	0.04	0.00	0.00	0.43	0.04	0.00	0.00	0.00	0.00
明確化要求に応える	2.35	0.08	1.97	0.25	1.61	0.13	0.85	0.08	0.00	0.00	0.00	0.00
会話を始める	0.00	0.00	0.66	0.08	0.00	0.00	0.43	0.04	0.00	0.00	0.00	0.00
トピックを変更する	0.00	0.00	1.97	0.25	0.00	0.00	2.99	0.29	0.00	0.00	0.00	0.00
会話を発展させる	0.00	0.00	4.92	0.63	6.99	0.54	6.41	0.63	5.36	0.25	0.00	0.00
決定する	9.41	0.33	0.00	0.00	13.44	1.04	7.69	0.75	0.00	0.00	0.00	0.00

注:1人あたりの平均が0.7以上を黒、0.35以上0.7未満をグレーで塗り潰している。

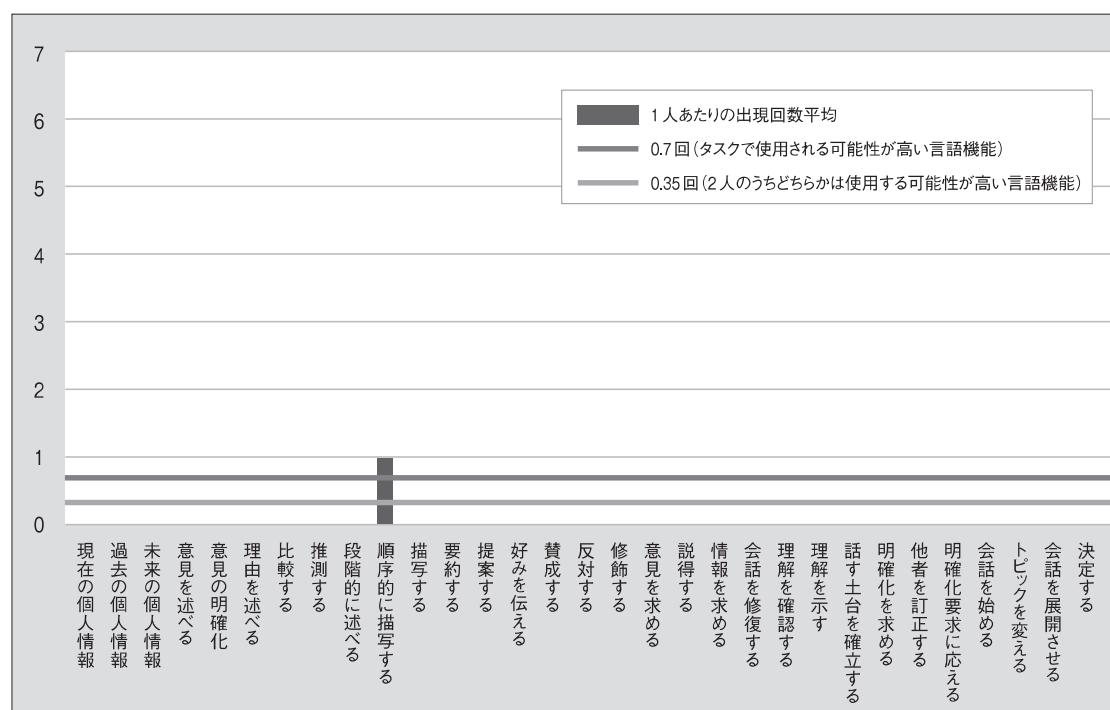
4.1 ナレーション

6種類のタスクタイプのうち、引き出される言語機能が最も少なかったタイプがナレーションである。ナレーションタイプのタスクとして用いられた4コマ漫画描写タスクの言語機能の分析結果を図2に示す。このタスクでは、参加者は4コマ漫画が渡され、30秒の準備時間の後、各コマの状況を描写することが求められた。結果として、時系列順に並ぶ4コマを描写するだけのタスクであったため、全ての参加者から引き出された言語機能は「順序的に描写する」だけであった。

この形式のタスクが用いられる代表的なテストとして実用英語技能検定の二次試験があげられ、2級では3コマ漫画を、準1級では4コマ漫画を描写するタスクをそれぞれ含んでいる。実際の二次試験では、タスクの前後に面接官との質疑応答が含まれている。そのため、テスト受験者は「順序的に描写する」以外の言語機能を他の場面で使用していることが考えられる。本研究では、参加者が行ったタスクは4コマ漫画の描写のみであったため、限られた言語機能のみが引き出される結

果となった。

ここで注意したいのは、Koizumi & Hirai (2012) が主張するように、引き出される言語機能が少ないタスクが必ずしも悪いタスクではなく、タスクの良し悪しは、あくまでタスクを用いて学習者にどのようなことをさせたいのか、つまりタスクを使用する目的によって決まるということである。他の種類のタスクとは異なる言語機能をもつ発話を引き出すことができるタスクは、多様な言語機能を引き出すタスクの組み合わせを考える場合には、重要な役割を担っていると考えられる。実際に、「順序的に描写する」という言語機能は他のタスクでは観察されなかつたことから、ナレーションタイプのタスクは他のタスクタイプとは異なる性質の発話を学習者から引き出しができると考えられる。



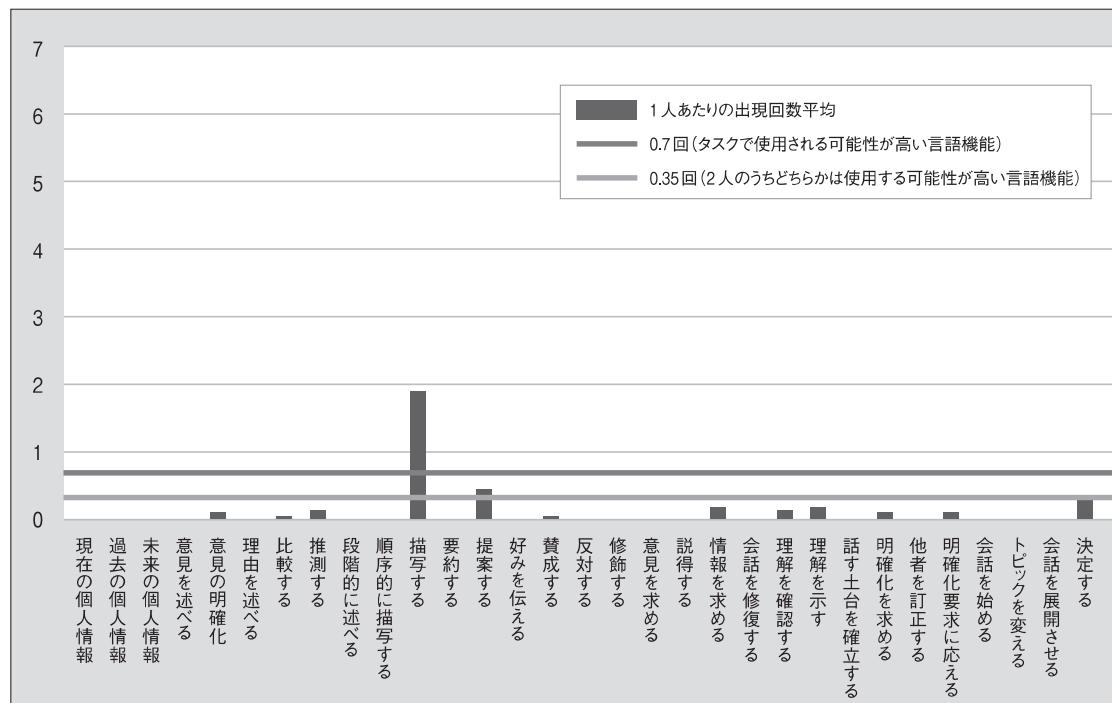
■図2: 4コマ漫画描写タスク中に用いられた言語機能

4.2 ジグソー

ナレーションタイプの後に引き出される言語機能が多かったタスクは、ジグソータイプとして使用した物語復元タスクである。図3に物語復元タスクの言語機能分析の結果を示す。このタスクは、ペアのそれぞれに対して4コマ漫画のうち2コマずつをランダムに配布し、相手に自分の絵を見せることなくもとの4コマ漫画のストーリーを完成させることを目的として行った。結果として、全ての参加者が使用することが期待される言語機能は、「描写する」のみであった。ペアのうちどちらか片方の参加者が使用することが期待される言語機能は「提案する」であり、合計2種類の言語機能がタスクによって引き出されたという結果になった。

本研究で使用した4コマ漫画は、実用英語技能検定準1級の二次試験で出題された4コマ漫画をもとにして作成しており、比較的内容やコマ間の関係が複雑ではなかったため、コマの順番を決定する際に参加者間のやりとりがそこまで必要ななかった可能性がある。その結果、ペアがそれぞれ持っているコマを「描写し」て、その後、話の筋が

通るようにそれぞれのコマの順番を「提案する」だけでタスクが終わってしまった可能性が高い。物語復元タスクは、市販の漫画の数コマを使用して作成することが可能なため、作成が容易なタスクの一つであり、コマ間の因果関係の複雑さや時間がどの程度離れているかといった要因で難易度が調整できるとされる(松村, 2017a)。コマ間の関係をより複雑にすることで、本研究で引き出すことができなかった「同意する」や「反対する」といった他の言語機能がタスク中に観察される可能性は十分に考えられる。



■図3：物語復元タスク中に用いられた言語機能

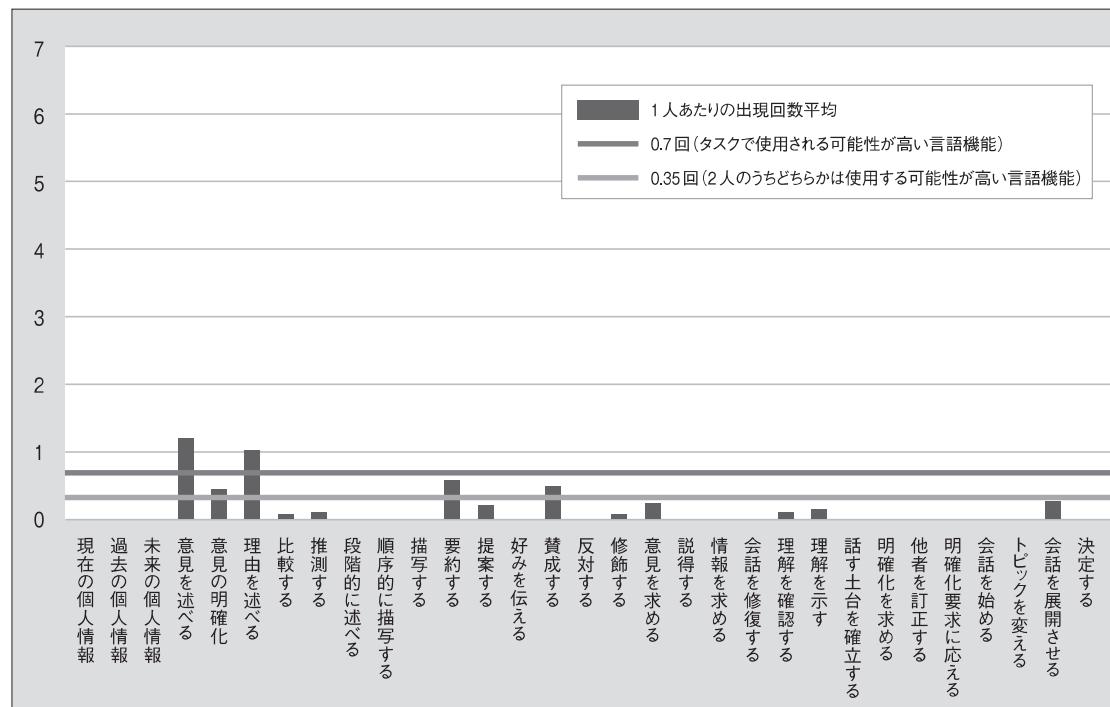
4.3 意見交換

ジグソータイプのタスクの後に多くの言語機能を引き出すことができたタスクタイプは意見交換タイプであった。図4に意見交換タスクにおける発話の言語機能分析の結果を示す。本研究では、与えられたトピックに対してできるだけ多くの意見を理由も含めて提案しあい、相手の意見に賛成または反対することまでが求められるタスクを使用した。言語機能分析の結果、全ての学習者が使用することが期待される言語機能は「意見を述べる」と「理由を述べる」の2種類であった。また、ペアのうちどちらか片方の参加者が使用することが期待される言語機能は、「意見の明確化」「要約する」「賛成する」の3種類であり、合計5種類の言語機能がこのタスクによって引き出されたという結果となった。

ナレーションタイプと情報ギャップタイプのタスクでは与えられた絵の描写を中心であるため、伝える内容を考える必要はなかった。一方で、意見交換タイプのタスクでは、伝える内容についても自分で考えなければならず、さらにそれを伝えなければならないため、引き出される言

語機能の傾向が異なる結果となった。「意見の明確化」では、例や言い換えを使用して、自分の意見をより理解しやすいものにしようとする発話が観察された。

しかしながら、「賛成する」の1人あたりの平均は0.46回であり、ペアあたりに換算すると0.92回であることから、2人のうちどちらかの参加者が意見を述べ、もう片方の参加者がその意見に同意するだけという会話になっている可能性が高い。また、「要約する」の1人あたりの平均は0.58回であり、ペアあたりに換算すると1.16回になることから、ペアの片方の参加者が「意見を述べ」、もう1人の参加者がそれに「賛成した」後、最終的に2人の意見を「要約して」いる可能性が高いことがわかる。これは、本研究で意見交換タスクを行う際に、2人の参加者が異なる立場から意見を述べることを要求していないため、意見が偏ってしまったことが理由として考えられる。「比較する」や「推測する」といった比較的難しい内容の会話をやって考えを深めたり、「反対する」といった言語機能を引き出してやりとりを活発にするためには、タスクの実施方法を工夫する必要がある。



■図4：意見交換タスク中に用いられた言語機能

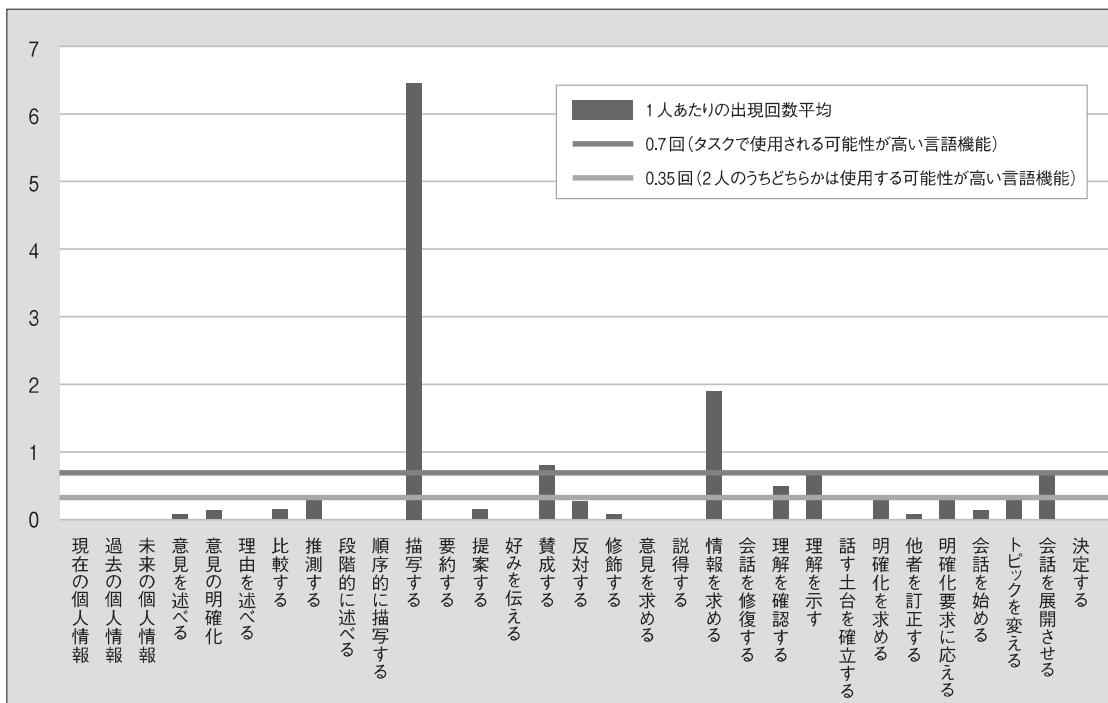
4.4 情報ギャップ

意見交換タイプの次に引き出される言語機能が多かったタスクタイプは、情報ギャップタイプであった。図5に情報ギャップタイプで使用した間違い探しタスクの言語機能分析の結果を示す。間違い探しタスクは、動物園の地図に描かれている9種類の動物のうち、ペア間で位置が異なる2種類の動物を探すタスクであった。言語機能分析の結果より、全ての参加者が使用することが期待される言語機能は「描写する」「情報を求める」「賛成する」の3種類であった。また、ペアのうちどちらか片方の参加者が使用することが期待される言語機能は「理解を確認する」「理解を示す」「会話を展開させる」の3種類という結果となった。合計で6種類の言語機能がタスクによって引き出されるという結果となり、このタスクは「描写」型のタスクでは最も多くの言語機能を引き出していた。

同じく「描写する」がタスクによって引き出されていた物語復元タスクとの違いとして、間違い探しタスクでは、情報を伝えるだけでなく「情報を求める」ことが必要だったことがあげられる。物語復元タスクでは手元にある2コマの説明をす

ればよかったですのに対して、間違い探しタスクでは9種類の動物の位置関係についてすべて把握することが求められる。そのため、間違い探しタスクにおいては、参加者は相手がもっている情報をうまく引き出しながら、他の動物との位置関係を何度も確認するために、「情報を求める」言語機能が必要だったと考えられる。実際に、タスク中の発話では、1度言及された動物について、後から動物の位置を再確認する会話が観察された。

さらに、これまでのタスクでは用いられることができなかった「会話を展開させる」が、1人あたり0.7回には届かないものの、1人あたり0.63回であり、ペアあたり1.26回タスク中に使用されていることがわかった。タスク中の会話では、例えば、“next to the horse, there is a pig. how about yours?”のような発話が観察され、同じトピックを継続させながらも、相手の発話を引き出すように会話を展開される場面が観察された。「会話を展開させる」は、後述する問題解決タイプと意思決定タイプにおいてもペアあたりの閾値である0.35を超えており、6種類のタスクのうち、多くの言語機能を引き出した3つのタスクに共通する特徴であった。



■図5：間違い探しタスク中に用いられた言語機能

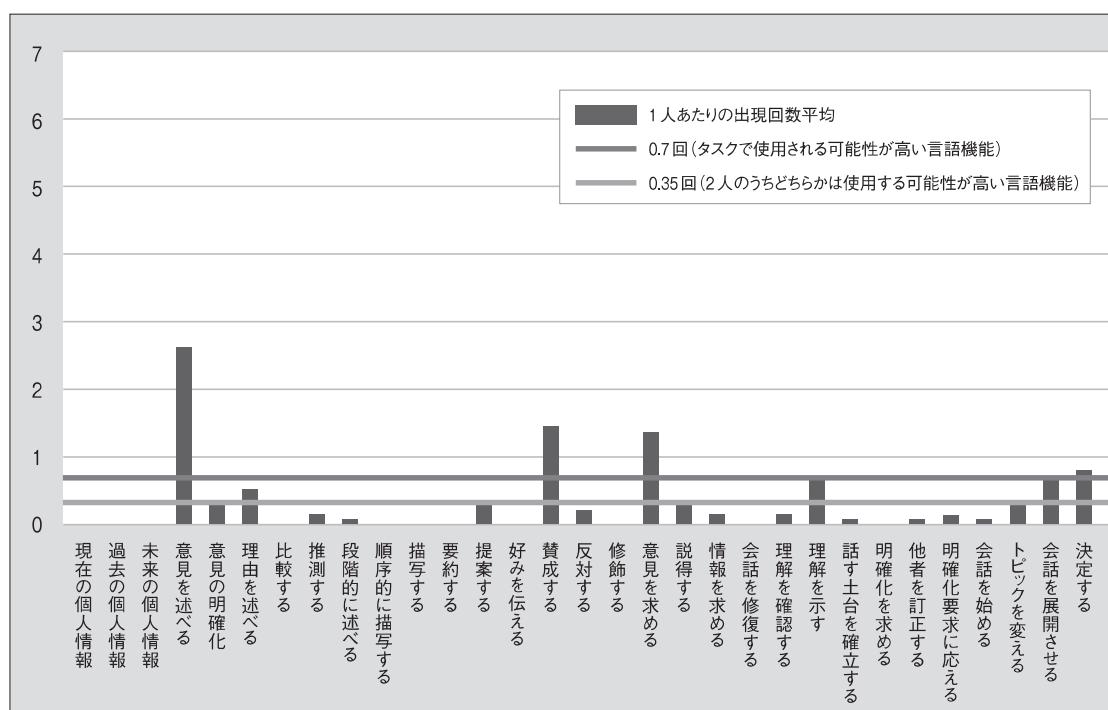
4.5 意思決定

情報ギャップタイプの次に多くの言語機能を引き出していたタスクタイプは意思決定タイプであった。図6に、意思決定タイプで用いられたランキング作成タスクの言語機能分析の結果を示す。ランキング作成タスクは、全部で15個あるアイテムのうち、ペアで上位5つを決めることが求められるタスクであった。分析の結果から、全ての参加者が使用することが期待される言語機能は、「意見を述べる」「賛成する」「意見を求める」「決定する」の4種類であり、ペアのうちどちらか片方の参加者が使用することが期待される言語機能は「理由を述べる」「理解を示す」「会話を展開させる」の3種類で合計7種類の言語機能がタスクによって引き出された結果となった。

意思決定タイプのタスクでは、意見交換タイプのタスクと同様に、参加者は「意見を述べる」ことが求められる。しかし、意見交換タスクと比べて検討するアイテムの数が増えているため、「意見を求める」という言語機能も使用されていたと考えられる。言及する必要があるものごとが増えた場合に相手の発言を促すような発話が増えた

ことは、ジグソーと情報ギャップの間にも同様に観察されていた。間違い探しでは、伝える内容が手元にあったために「情報を求める」が用いられていたが、ランキング作成タスクでは、タスクのゴールを達成するために手元の情報ではなくお互いの意見が必要だったため、「意見を求める」言語機能をもつ発話が用いられたと考えられる。

意思決定をする場では、根拠をもとに何かを決めることが求められるため、「決定する」と「理由を述べる」機能をもつ発話の使用が多くなることが事前に予想された。実際に、「決定する」については、1人あたりの平均が0.75回であり、ほとんどの参加者が「決定する」ための発話を行っていたことがわかった。しかしながら、「理由を述べる」については、1人あたりの平均が0.50回、ペアあたりの平均が1.00回であり、予想に反してペアのうちどちらか片方の参加者が使用する程度であった。これは、ランキングの上位に割り当てられるアイテムはそもそもペア間で考えが同じであり、理由まで述べる必要があったのは5位か6位かを決める最後の場面だけだった可能性が考えられる。



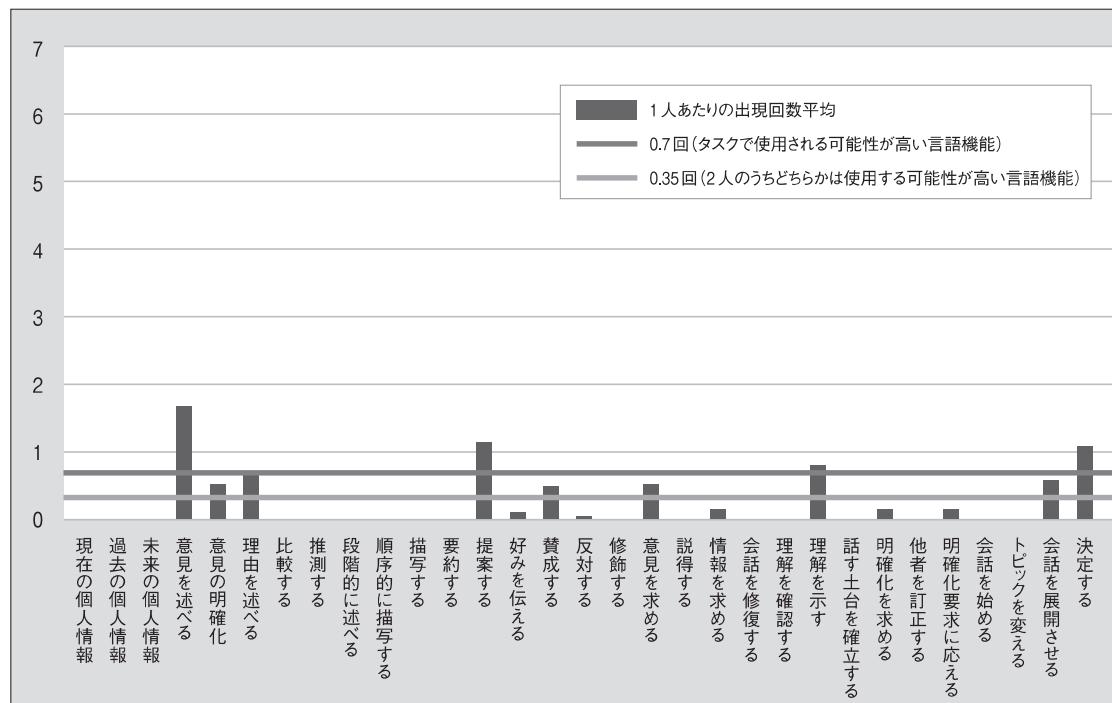
■図6：ランキング作成タスク中に用いられた言語機能

4.6 問題解決

6種類のタスクタイプのうち、最も多くの言語機能を引き出すことができていたタイプは、問題解決タイプであった。図7に問題解決タイプで使用した解決策提示タスクの言語機能分析の結果を示す。問題解決タイプのタスクは、与えられた問題に対して、各自1つ以上の解決策とその理由を考え、話し合いを通してペアで最も良い解決策を考えるタスクであった。結果として、全ての参加者が使用することが期待される言語機能は「意見を述べる」「提案する」「理解を示す」「決定する」の4種類であり、ペアのうちどちらか片方の参加者が使用することが期待される言語機能は「意見の明確化」「理由を述べる」「賛成する」「意見を求める」「会話を展開させる」の5種類であった。合計で9種類の言語機能がタスクによって引き出されたという結果となった。

解決策提示タスクは、他の「意見」型のタスクであるランキング作成タスクと意見交換タスクとは異なり、最終的に、ペアで1つの具体的な解決策を考え出す必要がある。そのため、「提案する」機能をもつ発話が1人あたりの平均が1.13回使用

され、「意見を述べる」とは別に全員が解決策を提案していることがわかった。これは、他の「意見」型のタスクには見られなかった特徴であると言える。それに加えて、ペアのうち1人は「理由を述べる」「賛成する」「意見を求める」「決定する」といった他の「意見」型タスクにおいても観察された言語機能を用いていた。このことから、本研究で用いられた問題解決タスクは、意見交換タイプと意思決定タイプを合わせたようなやりとりがタスク中に行われていたことが考えられる。



■図7：解決策提示タスク中に用いられた言語機能

以上のように、タスクのタイプによって、タスク中に引き出される言語機能は異なることがわかった。タスクのゴールを達成するために必要な言語機能が増えると、タスク中に必要な表現の幅が増え、やりとりが複雑になることやタスクのゴール達成が難しくなることが予想される。しかしながら、コミュニケーションタスクの利点である、多様な言語機能を引き出すということを最大限に活かすためには、学習者のレベルに合わせて、使用するタスクの組み合わせを変えていくことが有効であると考える。最低でも、「描写」型タスク(i.e., ナレーション, ジグソー, 情報ギャップ)と「意見」型タスク(i.e., 意見交換, 意思決定, 問題解決)を1つずつ使用すると良いだろう。

例えば、学習者の熟達度が低く、使用できる表現が制限されている場合には、求められる言語機能が少ない組み合わせであるナレーションタイプと意見交換タイプのタスクを組み合わせると良い。4コマや3コマから構成される漫画の絵を描写するタスクと漫画の内容に関連する事柄について意見を考えて述べさせるようなタスクなどが考えられるだろう。この組み合わせでは、「順序的に描写する」「意見を述べる」「理由を述べる」の3つの言語機能が最低限必要になる。実用英語技能検定2級および準1級の二次試験はこの組み合わせを採用していることもあり、学習者にとっては馴染みのあるタスクの組み合わせであると考えられる。

一方で、熟達度が高く、表現の幅が広い学習者に対しては、最も多くの言語機能を引き出すことが期待できる情報ギャップタイプと問題解決タイプの組み合わせが有効だろう。情報ギャップタイプと問題解決タイプを組み合わせて使用することで、「意見を述べる」「理由を述べる」「提案する」「理解を示す」「決定する」「描写する」「賛成する」「情報を求める」の8種類の言語機能が最低限使用されることになる。それに加え、ペアのうち片方の学習者が使用する可能性がある言語機能は「意見の明確化」「意見を求める」「理解を確認する」「会話を発展させる」の4種類であり、全てを合計すると、学習者は12種類の言語機能を使用してタスクを行うことになる。このように、学習者の熟達度や、授業等で扱った言葉の働きに合わせて、より多くの言語機能が引き出されるタスクタイプの組み

合わせを使用していくことが望ましい。

5 結論

本研究では、ジグソー、情報ギャップ、問題解決、意思決定、意見交換、ナレーションの6種類のタスクタイプについて、タスク中に引き出される言語機能の観点から各タスクタイプの特徴を調査した。結果として、6種類のタスクタイプは、引き出される言語機能という点ではそれぞれ異なる特徴を持っていることが明らかとなった。本研究の結果に基づく教室での指導や評価におけるタスクの使用に対する示唆は次の2つである。第一に、6種類のタスクタイプは「描写」型タスク(i.e., ナレーション, ジグソー, 情報ギャップ)か「意見」型タスク(i.e., 意見交換, 意思決定, 問題解決)によって学習者が使用する主要な言語機能が異なってくるため、組み合わせて使用することが望ましい。第二に、「描写」型のタスクはナレーション、ジグソー、情報ギャップの順に、「意見」型のタスクは意見交換、意思決定、問題解決の順に引き出される言語機能が多くなっていくため、学習者の熟達度や学習進度など、使用できる表現の幅に合わせて、なるべく幅広い言語機能を必要とするタスクを選択すると良いと考えられる。

今後の課題としては、タスクタイプが同じでも、細かい設定によって必要な言語機能が異なる可能性があるため、同じタスクタイプの中で異なるタスクを用いて調査を行う必要がある。また、本研究では、各タスクタイプにおいて引き出される言語機能は決して多いとは言えない結果となった。タスクは、言語使用以外のゴールを達成することが第一の目的である。そのため、学習者はタスク中に必要以上の発言をしない可能性が高い。コミュニケーションタスクは多様な言語機能をもつ発話を引き出すことが期待されるが、実際に引き出される言語機能は最低限であると考える必要がある。よって、多様な言語機能をもつ発話を引き出すためには、タスクを行う際に意図した言語機能が必要となるような工夫が必要である。今後は、各タスクタイプにおいて、どのような工夫をするとどの言語機能が引き出

されるようになるかについての詳細な検証も必要だと考えられる。また、今回は協力者の数が限られていたため、さらに多くの協力者が行ったとしても、同様な結果が得られるかどうか確認する必要があるだろう。

謝辞

本研究を実施する機会を与えてくださいました公益財団法人日本英語検定協会の皆様、選考委員の方々、特に、私の研究を担当し、ご指導くださった和田稔先生に心より御礼申し上げます。また、本研究の計画から親身にご指導をいただいた筑波大学大学院の平井明代先生に心より感謝申し上げます。最後に、調査にご協力いただきました、協力者の皆さんに深く御礼申し上げます。

参考文献

(*は引用文献)

- * Amagai, Y. (2018). Learning new words through information-gap tasks: evidence from utterances. *Annual Review of English Language Education (ARELE)*, 29, 225–240.
- * Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge University Press.
- * Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- * 藤田恵里子 (2020). 高校教科書おタスクの種類に関する考察. 江戸川大学紀要, 30, 241–250.
- * Green, A. (2012). *Language functions revisited: Theoretical and empirical bases for language construct definition across the ability range (Vol. 2)*. Cambridge University Press.
- * He, L., & Dai, Y. (2006). A corpus-based investigation into the validity of the CET-SET group discussion. *Language Testing*, 23, 370–401.
- * 小林潤子. (2019). CAN-DO リストを使った英語教育と学習者の意識. 関東甲信越英語教育学会誌, 33, 13–26.
- * Koizumi, R., & Hirai, A. (2012). Comparing story retelling speaking test with other peaking tests. *The Japan Association for Language Teachers (JALT) journal*, 34, 35–60.
- * 紺渡弘幸. (2015). 英語の授業における意見・考え方の表出を求める効果的な指導法: 問題解決型タスクの効果と課題. 中部地区英語教育学会紀要, 44, 163–168.
- * 松村昌紀. (2017a). 「タスクベースの英語指導: TBLTの理解と実践」東京:大修館書店.
- * 松村昌紀. (2017b). タスク・タイプの理論的基盤と学習者の言語使用. 中部地区英語教育学会紀要, 46, 55–62.
- * 文部科学省. (2013). 各中・高等学校の外国語教育における「CAN-DO リスト」の形での学習到達目標設定のための手引き. https://www.mext.go.jp/a_menu/kokusai/gaikokugo/_icsFiles/afieldfile/2013/05/08/1332306_4.pdf より (2020年5月20日閲覧)
- * Nation, I. S. P. (2013). *What should every EFL teacher know?*. Tokyo, Japan: Compass Publishing.
- * National Aeronautics and Space Act (n.d.). Nasa exercise: survival on the moon. Retrieved from <https://www.humber.ca/centreforteachingandlearning/assets/files/pdfs/MoonExercise.pdf> (2019年8月6日閲覧)
- * Nakatsuhara, F. (2014). A research report on the development of the test of English for academic purposes (TEAP) speaking test for Japanese university entrants-study1&2. https://www.eiken.or.jp/teap/group/pdf/teap_speaking_report1.pdf より (2020年5月20日閲覧)
- * O' Sullivan, B., Weir, C., and Saville, N. (2002). Using observation checklists to validate speaking-test tasks. *Language Testing*, 19, 33–56.
- * Pallotti, G. (2019). Assessing tasks: The case of interactional difficulty. *Applied Linguistics*, 40, 176–197.
- * Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, 43, 1–32.
- * Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey Bass
- * 柳川浩三. (2017). CLIL 型授業の実践-学習者はどう受けとめタスクは機能するか. 全国英語教育学会紀要, 28, 319–334.

資料1：Observation checklist (O'sullivan, Weir, & Saville, 2002)

<i>Informational functions</i>	
Providing personal information	<ul style="list-style-type: none">• Give information on present circumstances• Give information on past experiences• Give information on future plans
Expressing opinions	Express opinions
Elaborating	Elaborate on, or modify an opinion
Justifying opinions	Express reasons for assertions s/he had made
Comparing	Compare things/people/events
Speculating	Speculate
Staging	Separate out or interpret the parts of an issue
Describing	<ul style="list-style-type: none">• Describe a sequence of events• Describe a scene
Summarizing	Summarize what s/he has said
Suggesting	Suggest a particular idea
Expressing preferences	Express preferences
<i>Interactional functions</i>	
Agreeing	Agree with an assertion made by another speaker (apart from 'yeah' or nonverbal)
Disagreeing	Disagree with what another speaker says (apart from 'no' or nonverbal)
Modifying	Modify arguments or comments made by other speaker or by the test-taker in response to another speaker
Asking for opinions	Ask for opinions
Persuading	Attempt to persuade another person
Asking for information	Ask for information
Conversational repair	Repair breakdowns in interaction
Negotiating meaning	<ul style="list-style-type: none">• Check understanding• Indicate understanding of point made by partner• Establish common ground/purpose or strategy• Ask for clarification when an utterance is misheard or misinterpreted• Correct an utterance made by other speaker which is perceived to be incorrect or inaccurate• Respond to requests for clarification
<i>Managing interaction</i>	
Initiating	Start any interactions
Changing	Take the opportunity to change the topic
Reciprocating	Share the responsibility for developing the interaction
Deciding	Come to a decision

第32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

ループリックの事前提示が ライティングパフォーマンスに与える影響

研究者: 東京都／中央大学附属中学校・高等学校 教諭 岩本 祐樹

《研究助言者: 小泉 利恵》

概要

本研究は、パフォーマンス課題などを評価する際に使用されるループリックが、ライティング力の向上に有効かどうかを調査した。ループリックは評価基準が記述されているので、それを参照した上でライティングに取りかかれば、パフォーマンスの質が向上することが予想される。対象とした中学1年生(計60名)のうち1つのグループ(30名)には、ループリックを事前に提示した上で、ライティングを行ってもらい、フィードバックもループリックとスコアの両方を提示した。もう一方のグループ(30名)は、ループリックを提示されずにライティングをおこない、フィードバックとしてループリックに基づいて採点されたスコアのみを返却した。以上のライティングを3回実施し、その内容面と言語面のスコアを比較したが、グループ間で有意な差は見られなかった。以上の結果を踏まえると、ループリックの効果を高めるためには、ループリックの評価基準の明確化や、自己評価などのメタ認知活動や相互評価などを併用することが重要であるのかもしれない。

1

はじめに

現代社会における学校教育では、知識を習得するだけではなく、知識を使いこなしたり創造したりする力を育成すべきとの社会的 requirement が高まっている(石井, 2016)。知識を活用するといった場

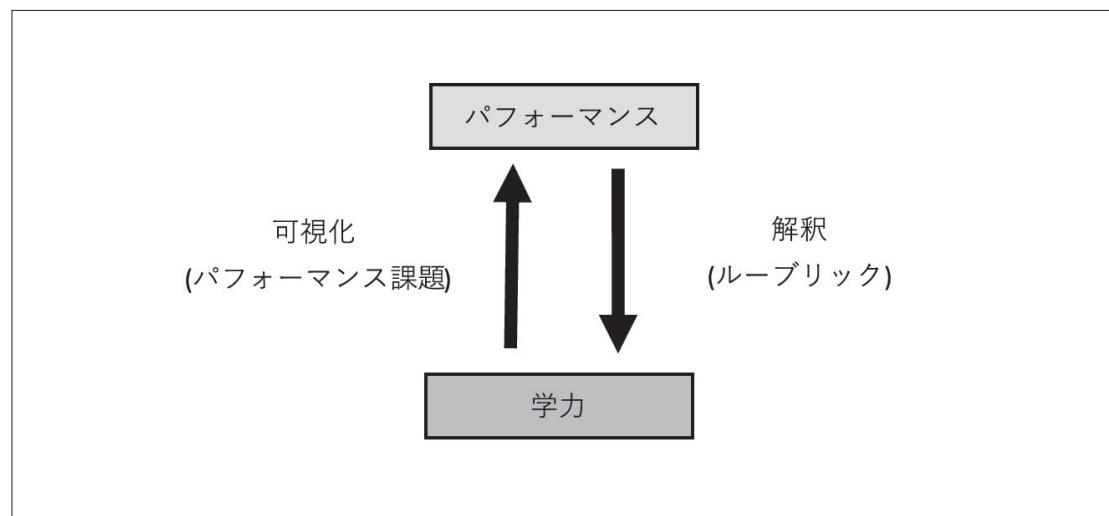
合、能力を階層的に捉えることが多い。例えば、石井(2015)は、教科内容に関する学びの深さは3つのレベルで捉えることができるとしている。1つ目は、知識・技能の習得状況を問う「知っている」レベル、2つ目は、概念の意味理解を問う「わかる」レベル、そして、3つ目は、現実世界の文脈における知識・技能の総合的な活用力を問う「使える」レベルである。

「使える」レベルの能力の重要性は、近年の知識と学びの「量」から「質」への転換を待たずしても、英語教育においては特に求められてきたことである。そして、このレベルの英語能力の育成を考えたときに不可欠であるのが、実際の英語使用場面を想定した指導や評価¹である。これを可能にするのが、パフォーマンス評価である。Wiggins & McTighe(2005)は、パフォーマンス課題を、「効果的に行動するために知識を活用する課題、あるいは、ある人の知識と熟達化を明らかにするような複雑な感情作品を実現する課題」(Wiggins & McTighe, 2005, 邦訳 p. 405)と定義している。言語教育におけるパフォーマンス課題は、プレゼンテーション、ディベート、スピーチ、演技といった実演を求めるものや、レポート、論文、ポスターなどの作品も含まれるであろう。そして、このパフォーマンス課題を評価する際に用いられるのが、ループリックである。松下(2007)は、学力には「見えやすい学力」と「見えにくい学力」があるとしているが、「使えるレベル」の学力は、「見えにくい学力」と言える。松下は、この「見えにくい学力」を可視化するのがパフォーマンス課題であり、ルー

ブリックを使うことでこのパフォーマンス課題から学力を「解釈」することが、パフォーマンス評価だとしている(図1)。

ループリックを使用した評価の実践は各教科で行われている。しかしながら、沖(2014)は、初等

中等教育におけるループリックの活用状況について、「毎回の授業や授業内の学習活動で効果的に提示あるいはフィードバックされているとは言い難い」と指摘しているように、その運用にはまだ問題点や実践的研究が不十分な点がある。



■図1: パフォーマンス評価の構図 (松下, 2007, p. 11)

また、言語教育におけるテスティングの観点からも、信頼性と妥当性のあるパフォーマンス評価の研究は行われている。教育現場では、ライティングやスピーキングのテストを実施する場合、教員がテーマを設定し、独自の評価尺度により判断を下す場合は多いが、藤森・小泉(2011)は、スピーキングテストの教員による全体評価と、生徒の発話を正確さ、流暢さ、複雑さの観点から客観的に分析した評価との関連を検証している。その結果、日本人教師と外国人英語講師の採点結果を総合したスコアは、正確さを厳しく客観的に評価した場合に有意な相関が見られた。

本実践研究では、日本語母語英語学習者のライティング課題に焦点を当て、評価におけるループリックの役割を検証する。後述するように、ループリックは、ペーパーテストでは測れない質的な評価をする道具であると同時に、教員が学習者に求める学習の到達レベルが記述されているため、形成的評価に使うことで教育効果をもたらし、指導と評価の一体化を具現化するツールとなりえる。ライティング課題は、語彙力、文法力、表現力、論理性、独創性などの総合的な知識が問われるパ

フォーマンス課題と言えるが、学習者がそのライティングをする際に、ループリックが提示され、求められていることを認識することができれば、直感的にはパフォーマンスの向上に有益であると推測できる。また、ループリックを用いたフィードバックにより、現在の到達レベルや足りない点を把握することができれば、次の課題にとりむる際に有効に働くと考えられる。本実践では、ライティング課題の実施の事前にループリックを提示することが、パフォーマンスの向上に有効かどうかを中学1年生のライティング活動を通して、実践的に検証することを目的としている。

2 先行研究

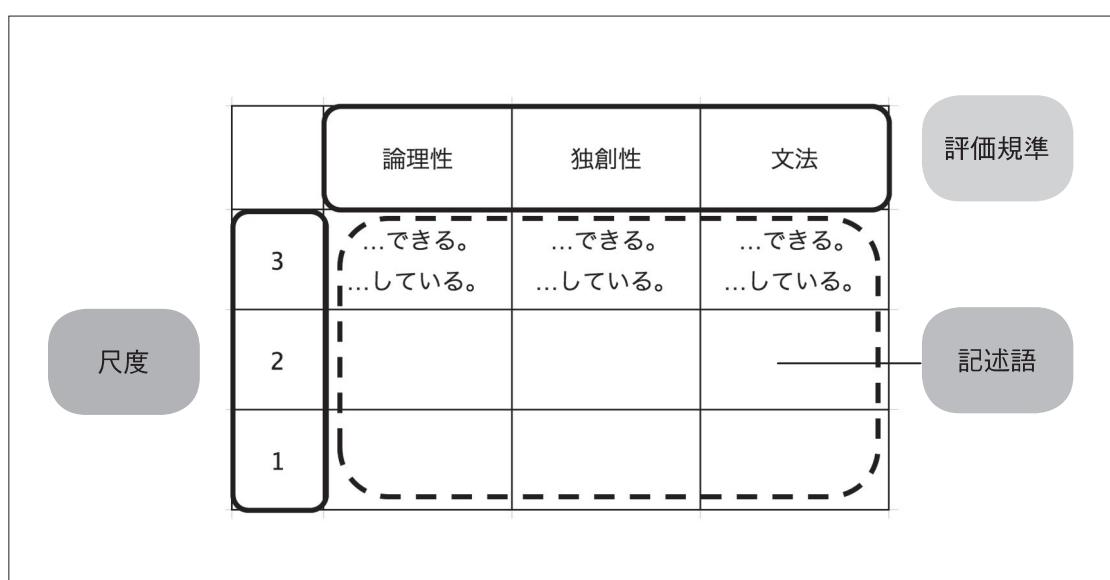
2.1 ループリックとは

ループリックは、主にパフォーマンス課題を評価する際に使われる道具のことであり、「成功の度合いを示す数値的な尺度(scale)と、それぞれ

の尺度に見られるパフォーマンスの特徴を示した記述語(descriptor)からなる評価指標」(田中, 2010, 48頁)と定義される。ルーブリックは、パフォーマンス課題のような質的評価をするための道具である。また、松下(2007)は、パフォーマンスから学力を解釈する際に、恣意的・独断的にならないようするためにルーブリックを使う必要があるとしている。また、沖(2014)は、ルーブリック評価は、テスト法や質問紙法を除いて、思考・判断や関心・意欲・態度、技能・表現を評価するすべての評価方法に適しているということができる、と指摘している。

ルーブリックには、いくつかの種類がある。代表的なものが「分析的ルーブリック」と「全般的

ルーブリック」の2つである。前者は、ある課題を評価する際に、複数の評価規準²を設定し、それについてレベル分けをして(尺度)、評価基準(判断基準)を記述しているものである。後者は、全ての規準を1つにして、レベルを設定し、評価基準を記述しているものである。図2は、分析的ルーブリックの概要を表している。また、ルーブリックの各レベルに対応する典型的なパフォーマンスの事例をアンカー作品と呼ぶ。アンカー作品をルーブリックに添付することで、学習者は課題が求めていることをより明確に理解することができる。このように、評価規準と評価基準を明示することにより、学習者は自身の到達度や目指すべきレベルとの差を知ることができる。



■図2: 分析的ルーブリックのイメージ

ルーブリックは質的なパフォーマンス課題を恣意的・独断的にならないよう評価するために使用するのだが、当然ながら評価者の主觀がまったくなくなるわけではない。同じルーブリックを用いて評価しても、評価する人によって結果が異なることはもちろんあり得る。ルーブリックによる評価が恣意的になるのを防ぐためには、主觀のつき合わせや調整によって間主觀性を担保することが求められる(松下, 2012)。そして、間主觀性を担保するために用いられるのが、モデレーションである。モデレーションは、評価過程を統

一する方法と評価結果を統一する方法の2つがある。前者は、評価基準を作成する際に、複数の評価者で作品例を評価し、その結果を比較・検討し、共同でルーブリックの作成・修正を行い、評価規準と評価基準についての共通理解を図り、評価結果を統一することである。後者は、評価の結果を複数の評価者間で調整することである。

パフォーマンス評価の際に使用されるルーブリックには、様々な利点があることが指摘されている。Stevens & Levi(2005)は、ルーブリックの利点として、①タイミングの良いフィードバッ

ク, ②学生による詳細なフィードバックの利用, ③批判的思考力のトレーニング, ④他者とのコミュニケーションの活性化, ⑤教員の教育技法の向上, ⑥平等な学習環境作り, の6点を挙げている。この中で最も重要であると考えられる要素が, ①と②のフィードバックの側面である。パフォーマンス課題を評価する場合、評価が曖昧になつたり、コメントなどに時間がかかったりする場合が多くある。ループリックには、評価基準、到達の度合い、不足している部分がすでに記述されているため、教員は、多くの時間を割くことなく、タイミングの良いフィードバックをすることができる。必要に応じて、コメントを追加することも可能である。一方、生徒も比較的迅速なフィードバックを期待しており、学習を促進するためには、適切なタイミングでフィードバックをすることが必要となる。ループリックを使えば、時間がかかる質的な評価を多くの時間を割くことなくおこなうことができる。フィードバックにおいては、多くの情報が含まれていることと、どのような場合に最高の評価になるかが記載されていることが有効であるとされており、ループリックはこの両方を満たしていると言える(Stevens & Levi, 2005)。また、この他にも、後述するように形成的に使うことで多くの利点がある。

2.2 形成的評価としてのループリック

ここまで述べてきたように、ループリックは、パフォーマンス課題のような質的なものを評価し、適切なフィードバックを可能にする道具であり、質を量に変換する装置という一面を持つ(松下, 2012)。しかしながら、安藤(2014)は、成績評価のための採点ループリックだけではなく、形成的評価の観点からループリックを位置付けなければ、ループリックの教育効果が十分に発揮されないとしている。つまり、必ずしも点数化する必要はない、学習の最後ではなく、途上でループリックを学習者と共有し、期待する事柄を周知徹底することの必要性に言及している。Stiggins(2001)は、形成的に生徒中心の評価としてループリックを使うことの効果を指摘している。

Panadro & Jonsson(2013)は、小学校から大学までの21本のループリック研究を分析し、

ループリックの形成的な評価がパフォーマンスの向上に寄与する要素として、①評価の透明性、②フィードバックの省察、③自己効力感の向上、④不安の低減、の4点を挙げている。また、Panadro & Jonssonは、このパフォーマンスの向上に寄与する要素を効果的にする要因として、自己評価とメタ認知活動を挙げている。特に、後述の実証研究で紹介するように、ループリックを用いた自己評価と相互評価の有効性については多くの研究がある。山田・森・毛利・岩崎・田中(2015)は、ループリックを用いて、生徒を評価活動に参加させることで、生徒は求められる結果がどの程度達成される途上にあるか、またどの程度達成されたのかについて、より深い認識を持つことができるとしている。このように、評価を学びのプロセスの中に取り組み、評価活動に参加させることも、ループリックを用いること可能になる。

2.3 実証研究

前節で示したループリックの理論的な枠組みに対し、その効果を実証的に調査した研究がある。Arter, Spandel, Culham & Polland(1994)は、「書き方の特性(writing traits)」³と呼ばれるループリックを書き方の指導に導入する効果を検証した。書き方の特性は、アイデア、構成、ボイス、言葉の選択、文章の流暢さ、句読点等の取り決めの6つに、オプションとして、プレゼンテーションを含んだループリックである。Arterらは、小学5年生を対象に、このループリックを用いて、自身の作品の自己評価に加え、サンプルや他の生徒の作品を評価したり、改善策を考えたり、また教員によるそれぞれの尺度についての指導などの授業を行なつたりした。実験群(67名)と、ループリックを用いないで通常の書き方の指導をした統制群(65名)の事前テストと事後テストの作文を比較した。その結果、書き方の特性のうち、実験群の方がアイデアの観点において有意に統制群より優れていた。以上の結果を踏まえて、Alterらは、自己評価と評価基準を指導することの重要性を指摘している。

ライティングの評価からは離れるが、鈴木(2011)は、数学の授業とテストの評価にループ

リック評価を用いることで、ルーブリックがどのように学習者のテストに対する認識に影響を与えるかを検証した。鈴木は、評価の目的や基準が実施者と受け手との間で合意がなされている場合、その評価の仕方(インフォームドアセスマント)は、学習者のテストに対する認識(テスト観)をポジティブなものにするとしている。鈴木が挙げているテスト観とは、テストは自分の学習改善に活用するものであるという「改善」、自発的に学習を行うためのものであるという「誘導」、他者と比較するためのものだという「比較」、勉強を強制させるためのものだという「強制」の4つの因子で構成されるとしている。そして、ルーブリックは高い水準のインフォームドアセスマントとなりうると指摘し、ルーブリックの提示が学習者のテスト観に与える影響について、中学2年生の5日間の数学の授業において検証した。その結果、「改善」のテスト観において、ルーブリックを提示されたグループは学習改善に活用するための目的や役割を有意に強く認識していた。また、5日間のプログラムの最後に実施された総合テストでは、5日間の授業で扱った問題では、ルーブリックを提示された群は、有意に高い結果を示したが、授業で扱ったテーマとは無関係な発展問題においては、有意な差は見られなかった。

寺嶋、林(2006)は、大学の演習型の授業において、ルーブリックを事前提示し、学習の途中や終了時に学習者自身がそれをもとに自己評価する授業を実施した。受講者へ行った質問紙の結果から、ルーブリックをもとに自己評価をする事で、学習目標を意識化し、それに従って学習を進め、その学習における成果や課題を把握できるといった、学習一般への効果もあるということが示唆されている。

一方で、ルーブリックの提示がライティングに有意な効果を示さなかった研究もある。例えば、Goodrich Andrade (2001) は、形成的評価としてルーブリックを提示する事が、学習者の良いライティングに関する知識に影響を与えるかどうか、また実際にライティングのスコアに影響を与えるかどうかを検証した。この実験において、中学2年生を対象にルーブリックを提示するグループと提示しないグループに3種類のライティング(説得型エッセイと旅行記と自分の挑戦、または歴史的物語)を課し、そのスコアを比較した。また

最後に、「優れているライティングと良いライティングの違いを教員がどのように判断していると思うか」という記述式の質問をして、良いライティングに対する知識の度合いを分析した。その結果、ルーブリックを提示されたグループの方が、ライティングの評価基準に関する深い知識・理解は見られたものの、実際のライティングのスコアの比較では、3つのうち1つのライティング(旅行記)でしか有意な差は見られず、良いライティングに関する理解が実際のライティングのスコアに反映されることには、より難しいことであるとしている。

また、Goodrich Andrade & Boulay (2003) は、中学1年生と2年生に対し、ライティング(歴史的事語と物語を読んでの感想)における自己評価がパフォーマンスに及ぼす影響を調査した。実験群と統制群の両方にルーブリックを提示し、ライティングを1ヶ月の間をあけて、2回実施させた。その中で、実験群は、ルーブリックに沿って自己評価をする授業を実施した上で推敲する時間を設けたが、統制群に対しては単に推敲の時間を設けるのみにした。しかしながら、最終的な両者のライティングのスコアに有意な差は見られなかつた。Goodrich Andrade & Boulay は、自己評価の効果はあったかもしれないが、ルーブリックの提示の効果が大きく、自己評価の効果が顕著にならなかつたのかもしれない、としている。

英語教育においてもルーブリックを活用する実践は多い。例えば、鶴田(2018) は、高校生のスピーキングの帯活動にルーブリックを用いて、自己評価させる活動を行なった。スピーキングを「発表」と「やりとり」の2つにわけ、各6回ずつ活動をし、回ごとにルーブリックから焦点を当てる項目を1つ選び、その項目に意識して活動を行うように指示し、計12回の活動の中で段階的な指導を実践した。その結果、回を追うごとに自己評価は高くなり、ルーブリックを自己評価と共に継続的かつ段階的に使用することで、自己肯定感を高めるスピーキング活動が可能になるとしている。

また、松井(2015) は、高校生のライティング活動において、ルーブリックの作成を生徒と共同で行い、また自己評価をさせる実践を行った。その実践から、教員が作成したルーブリックを使うよりも、生徒が能動的に作成に関わったルーブリックを用いた方が、ライティング能力の育成に有効

であると指摘している。また、生徒と作成したルーブリックで自己評価を行うことは、教員の評価とのずれが少なくなり、自己評価力が高まるとしている。

以上の先行研究を踏まえると以下のことが言える。

1. ルーブリックを事前提示したり、ルーブリックを用いて書き方の指導をすることで、学習者は目指すべきものを理解したり、良い作品についての理解を深めることができるので、パフォーマンスが向上するかもしれない。
2. 自分のパフォーマンス課題を自己評価したり、他者のパフォーマンス課題を評価したりすることで、モチベーションやテストに対する意識、また自己調整力、自己評価力などのメタ認知能力の向上に影響を与えるかもしれない。
3. 2の結果として、パフォーマンスが向上するかもしれない。

このように、ルーブリックの提示やルーブリックを使ってのフィードバック、また、ルーブリックを用いての自己評価が、学習者のメタ認知能力の向上に寄与しそうであることは、学習者に対する質問紙などの結果から示唆されている。しかし、実際のパフォーマンスの向上に影響があるのか、またあるとしたらどのような面に効果があるのかは明確ではなく、ルーブリックの有無で、パフォーマンスのスコアを比較した研究は多くない。また、中学生のライティング活動においての検証も少ない。そこで、本実践研究は、上記1の点に焦点を当て、日本語を母語とする英語学習者である中学1年生を対象にしたライティング課題を取り上げ、ルーブリックの事前提示やフィードバックがパフォーマンスの質の向上に有効であるかどうかを検証する。

3 実践方法

3.1 目的

本実践の目的は、ライティング活動をループ

リック評価する際に、ルーブリックを学習者に事前提示することが、ライティングのパフォーマンスにどのような影響を与えるかを調査することである。英語のライティングは、文法知識、語彙力、論理性や独創性などの総合的な力が問われるパフォーマンス課題である。先行研究によれば、ルーブリックを事前に提示することは評価の透明性をもたらし、その課題が求めていることが学習者に明確になることで、パフォーマンスの質も向上することが期待できる。また、ライティングを複数回実施し、ルーブリックを伴うフィードバックをした上で、新たに同じ種類のライティングを行なえば、その効果は事前提示の効果に加えて大きくなると考えられる。しかし、事前提示に対しては、事後の質問紙などにより学習者から肯定的な認識はあるものの、その効果が実践において調査された研究は多くはない。そこで、本実践は日本語母語英語学習者である中学1年生を対象にライティング課題を実施し、その論理性や一貫性などの内容面と文法や語彙などの言語面の2つの側面⁴におけるルーブリックの事前提示とフィードバックの効果を検証する。以下の研究課題を設置し、3回のライティングの授業実践と事後のアンケートを実施する。

RQ1 ライティング課題において、ルーブリックを事前に提示した場合、そうでない場合と比較し内容面においてパフォーマンスが向上するか

RQ2 ライティング課題において、ルーブリックを事前に提示した場合、そうでない場合と比較し言語面においてパフォーマンスが向上するか

RQ3 ライティング課題において、ルーブリックを使用した場合、そうでない場合と比較し、課題に対する目標意識や、その後の学習への意欲が高くなる傾向が見られるか

3.2 参加者

本実践には、中央大学附属中学校の1年生の2クラスの69名が参加した。本実践は、ループリックの事前提示の効果を調査することが目的であるため、34名からなる1つのクラスにはループリックを提示し、フィードバックもループリックによって実施した(実験群)。そして、35名からなるもう一方のクラスは、ループリック自体を使用しない条件(統制群)のもとライティング課題を実施した。

クラス間の英語運用能力の均等性を可能な限り確保するために、英検の取得級を元に事前に分析の対象となる参加者を決定した。その結果、どちらのクラス(グループ)も、英検3級取得者が12名、4級取得者が18名となり、それぞれ計30名を分析の対象とした。

3.3 実践と調査の手順

3.3.1 ライティング課題のテーマ

本実践は、上記のように、ループリックを提示する実験群と、ループリックを全く提示しない統制群に分けて実施した。どちらのグループにも、3回のライティングをおよそ3週間の間隔で課した。ライティングの内容は、①1年間でどの月が好きですか、②日本のどの都市が好きですか、③どのテレビ番組が好きですか、の3種類を用意した。それぞれのライティング課題は難易度も異なり、また順序による効果があることも予想されることから、表1に示すようにそれぞれのグループを10名(英検3級取得者4名、英検4級取得者6名)ずつの3つのグループに分け、カウンターバランスをとった。

3.3.2 ループリック

本実践では、ライティング課題を、内容面と言語面の2つの側面で評価するために、主に論理的な主張ができているかを評価する尺度(内容面)と文法や語彙の使用について評価する尺度(言語面)の2つの尺度が含まれるループリックを使用した(表2)。このループリックは、国際バカロアの1つの科目であるEnglish Bのライティング評価で使われているものを中学1年生にも理解できる表現に訳して、作成したものである。

3.3.3 実施手順

表3は、本実践の概要を示している。本実践は、実践実施者が勤務する中学校での週5時間ある英語の授業内に実施した。実験群には、1回目のライティングをする前に、ループリックを提示し、ループリックの概要説明と、それに従って採点がなされることを説明した。また、内容面のループリックにのみ、評価5、評価4、評価3、評価2、評価1に対応するライティングの例(アンカー)を添付した(資料1)。アンカーは、「どの教科が好きですか」というテーマのもとにそれぞれの尺度に当たる答案を作成し、生徒にその理由とともに提示した。特に、評価5に相当する答案の例に関しては、論理性と一貫性があることをループリックの記述とアンカーを参照しながら説明した。さらに、アンカーとは別に、評価2に相当する答案をもう1つ用意し、評価2であることを言わずに、それがどの評価に該当するかを生徒自身に考えてもらい、その後評価2であることとその理由を提示した。そして、それらの全ての説明が終了した後に、1回目のライティングに移行した。ライティングの際は、教材として生徒が共通に使用している紙の和英・英和辞書の使用を許可した。

■表1: カウンターバランスによるライティング課題の振り分け

	実験群(30名)			統制群(30名)		
	グループ1	グループ2	グループ3	グループ1	グループ2	グループ3
第1回	月	都市	TV番組	月	都市	TV番組
第2回	TV番組	月	都市	TV番組	月	都市
第3回	都市	TV番組	月	都市	TV番組	月

一方で、統制群に対しては、何も提示せずにライティング課題を実施した。統制群においても紙の辞書の使用を認めた。ライティングはどちらのグループも20分間で行なった。

2回目のライティングは、1回目からおよそ3週間後に実施された。実験群においては、ルーブリックに採点されたスコアを表記し、答案と共に返却した(採点については後述)。それ以外は、評価やルーブリックの説明はせず、2回目のライティン

グを行なった。統制群については、採点者が実験群で使用しているルーブリックを使い評価を行ない、そのスコアのみを生徒に返却した。よって、生徒は、内容面と言語面において、それぞれ0点から5点の点数がフィードバックされるだけであった。そして、その後、2回目のライティングに入った。

3回目のライティングも2回目からおよそ3週間に実施した。3回目はどちらの群も2回目と同様の手続きで実施したが、ライティング後、それ

■表2: 本実践で使用したルーブリック(International Baccalaureate, 2011, Language B guideより抜粋し筆者が修正。)

	内容面	言語面
5	<p>理由は説得力がある。説明の流れも非常に良く、わかりやすい。</p> <ul style="list-style-type: none"> 理由は、具体的に非常に良く表現されており、妥当である。 また、興味深いものである。 ライティングの構成は、首尾一貫してわかりやすく、整理されている。 	<p>文法・語彙、文の作り方が非常に良い。</p> <ul style="list-style-type: none"> 語彙が豊富で、正確かつ効果的に使われている。 主語・動詞のある文をしっかりと作ることができている。 間違いがほとんどない。
4	<p>理由は明確で十分な説得力がある。説明の流れも良く、わかりやすい。</p> <ul style="list-style-type: none"> 理由は、十分に表現されており、妥当である。 ライティングの構成は、わかりやすく、整理されている。 	<p>文法・語彙、文の作り方が良い。</p> <ul style="list-style-type: none"> 語彙が豊富で、正確に使われている。 主語・動詞のある文をしっかりと作ることができている。 細かい間違いは多少あるが、意味に関わる間違いはほとんどない。
3	<p>理由はある程度明確で、部分的に説得力がある。</p> <p>説明の流れは、良い箇所がある。</p> <ul style="list-style-type: none"> 理由は、おおむね妥当である。 ライティングの構成は、明確である。 	<p>文法・語彙、文の作り方は適切であるが、誤りがいくつかある。</p> <ul style="list-style-type: none"> 最低限の語彙を用い表現できている。 主語・動詞のある文をおおむね作ることができている。 意味に関わるものも含め、いくつか間違いがある。
2	<p>理由はほとんど明確ではなく、説得力もほとんどない。</p> <p>説明の流れもあまり良くない。</p> <ul style="list-style-type: none"> 理由は、妥当である箇所がある。 ライティングの構成は、明確である部分もある。 	<p>文法・語彙、文の作り方はおおむね正しいが、誤りが多くある。</p> <ul style="list-style-type: none"> 語彙の種類が限られている。 主語や動詞のない文が多く見られる。 意味に関わるものも含め、間違いが多くある。
1	<p>理由は明確ではなく、説得力がない。説明の流れも悪い。</p> <ul style="list-style-type: none"> 理由は、見当外である。 ライティングの構成が曖昧で、わかりにくい。 	<p>文法・語彙、文の作り方が限られており、誤りが多くある。</p> <ul style="list-style-type: none"> 語彙の種類がかなり限られている。 主語や動詞がある文がほとんどない。 意味に関わるものも含め、間違いがかなり多くある。
0	上記に記されている基準に達していない。	上記に記されている基準に達していない。

そのグループに今回のライティング課題に関するアンケートを行った(後述)。

このような実験計画であるので、ループリックの事前提示による純粋な効果は1回目のスコアを比較することで行う。2回目以降は、その前のライティングのフィードバックがなされるため、ループリックの事前提示による効果とフィードバックの効果が混合したものになる。したがって、2回目と3回目のライティングにおいては、その前の自分自身のライティングのフィードバックも含めた上でのループリックの事前提示効果と解釈し、その効果を検証した。なお、ループリックによる評価を提示せずに、3回書いてもらうことも考えたが、採点後に点数の返却は通常行なうため、実際の指導に近い形の実験計画にした。また、この計画では、1回目では、実験群におけるループリックの提示や例示、解説の有無が統制群と異なる点であり、2回目、3回目では、ループリックの提示の有無がグループでの相違点となる。しかし、2回目、3回目での実験群へのループリック提示は、1回目での例示・解説を含むループリックの事前提示を思い出してもらう程度とも考えられる。

3.3.4 採点とフィードバック

回収された答案は、中央大学附属中学校・高等

学校に所属する英語科教員2名に採点を依頼した。採点者は、採点のブレが大きくならないように、筆者と共に、ループリックと生徒にも提示したアンケートをもとに、採点の基準について話し合った。採点する際には、採点者同士のすり合わせは行わず、それぞれがループリックをもとに評価を行なった。フィードバックには、2名の採点者が評価したスコアの平均をとり、小数点以下は切り捨てた数値を採用した。これは、例えば評価が3.5になつた場合、その内容は、ループリックの評価4には達していないく、3の範疇の中にあると判断したためである。統制群においても、同じループリックをもとに採点を行なった。なお、採点者には、実践の内容は明らかにせず、実験群と統制群が存在することは知らされず、60名分の答案をランダムにしたのちに、採点をしてもらった。

3.3.5 事後アンケート

表4に示すのが、3回目のライティング課題が終わった後に、実験群と統制群に行なったアンケートである。質問の(6)～(16)は、両グループに共通した質問項目であり、ライティング中に各要素(一貫性、構成、文法、語彙、評価へのモチベーション)を意識したかどうか、フィードバックに対する感想、今回のライティングは英語学習に有益であったか、また今後のモチベーション

■表3: 本実践の概要

	グループ 内容	実験群	統制群
第1回	ライティング①の実施	<ul style="list-style-type: none"> ・ループリックの提示 ・ループリックの説明 ・各評価尺度の例を提示、その理由を説明 ・ライティング①を実施 	<ul style="list-style-type: none"> ・ライティング①を実施 ※特に指示は無し
第2回	ライティング①の フィードバック	・ループリック評価を返却	・ループリックによる評価の点数のみを返却
	ライティング②の実施	・ループリックを提示し、ライティング②を実施	・ライティング②を実施
第3回	ライティング②の フィードバック	・ループリック評価を返却	・ループリックによる評価の点数のみを返却
	ライティング③の実施	<ul style="list-style-type: none"> ・ループリックを提示し、ライティング③を実施 ・アンケートを実施 	<ul style="list-style-type: none"> ・ライティング③を実施 ・アンケートを実施
第4回	ライティング③の フィードバック	・ループリック評価を返却	・ループリックによる評価の点数のみを返却

ンとなったかの主に3要素から成る。また、実験群のみに対して、ループリックの理解度と使用を確認するために、(1)～(5)の質問を聞いた。また、どちらのグループに対しても、最後に自由記述欄を設けた。

このようなアンケートを実施することで、両グループの生徒が今回のライティング学習に対して持った認識において、グループ間にどのような傾向があるのかを調査した。

■表4: 事後アンケートの項目

(1) 授業中に行ったループリックの説明は理解できましたか。
(2) その際に、例として提示されたWritingは、なぜその評価になるのか理解できましたか。
(3) ループリックの評価の記述はわかりやすかったですか。
(4) ループリックはWritingを書く上で、役に立ちましたか。
(5) Writing中に、ループリックの評価基準を確認したり、意識したりしましたか。
(6) Writing中に、内容が筋の通ったWritingにしようと意識しましたか。
(7) Writing中に、全体の構成を良くしようと努力しましたか。
(8) Writing中に、文法のミスを少なくしようと努力しましたか。
(9) Writingの中で、いろいろな語彙を使おうと努力しましたか。
(10) Writingをしている際に、良い評価を取りたいと思いましたか。
(11) 返却された答案が、その評価になる理由が理解できましたか。
(12) 返却された1回目の評価は、2回目のWritingに、役に立ちましたか。
(13) 返却された2回目の評価は、3回目のWritingに、役に立ちましたか。
(14) 計3回のWritingは英語の勉強になったと思いますか。
(15) 3回分のWritingを書いたことで、今後もっと英語の勉強をしようと思いましたか。
(16) 今回のようなWritingを繰り返し行ったら、Writingが向上すると思いますか。
※生徒は、1. 全然そうではない 2. ほとんどそうではない 3. どちらとも言えない 4. 少しそうだ 5. とてもそうだ の5段階で回答。 ※(1)～(5)の質問は実験群のみ。(6)～(16)の質問は2群共通。

4 検証方法と結果

ライティングのスコアは、分析においては2人の評価者が採点したものとの平均をとり、その数値をスコアとして採用した。したがって、生徒にフィードバックした評価とは異なり、小数点以下も含む数値で分析にかけた。また、本実践で行う統計処理は、全て有意水準を5%に設定し、分析を行った。

2人の採点者の評価に一貫性があるかどうかを相関関係の度合いで調査した。その結果、1回目、2回目、3回目の内容面のスコアはそれぞれ、 $r = .761$ ($p < .001$), $r = .671$ ($p < .001$), $r = .536$ ($p < .01$) となり、1回目は強い相関、2回目、3回目は、中程度の相関が確認された。また、一致度を検定するカッパ係数は、 $k = .682$, $k = .460$, $k = .467$ となり、1回目がかなりの一一致、2回目、3回目で中程度の一一致となった。

言語面のスコアは、1回目、2回目、3回目でそれぞれ、 $r = .499$ ($p < .001$), $r = .693$ ($p < .001$), $r = .236$ ($p = .208$) で、1回目、2回目で中程度の相関が見られたが、3回目は弱い相関はあるものの、有意な相関がなかった。カッパ係数は、 $k = .300$, $k = .557$, $k = .194$ であり、1回目がまずまずの一一致、2回目が中程度の一一致が確認されたが、3回目は、わずかに一致するにとどまった。

言語面の3回目の相関には課題があるものの、それ以外は、中程度の相関以上が確認できたので、これらのスコアを用いて、それぞれのライティングのスコアの平均を比較した。

なお、今後の実践や研究への参考のために、内容面と言語面が共に評価5の生徒のライティングと、共に評価3の生徒のライティングを資料として添付する(資料2)。

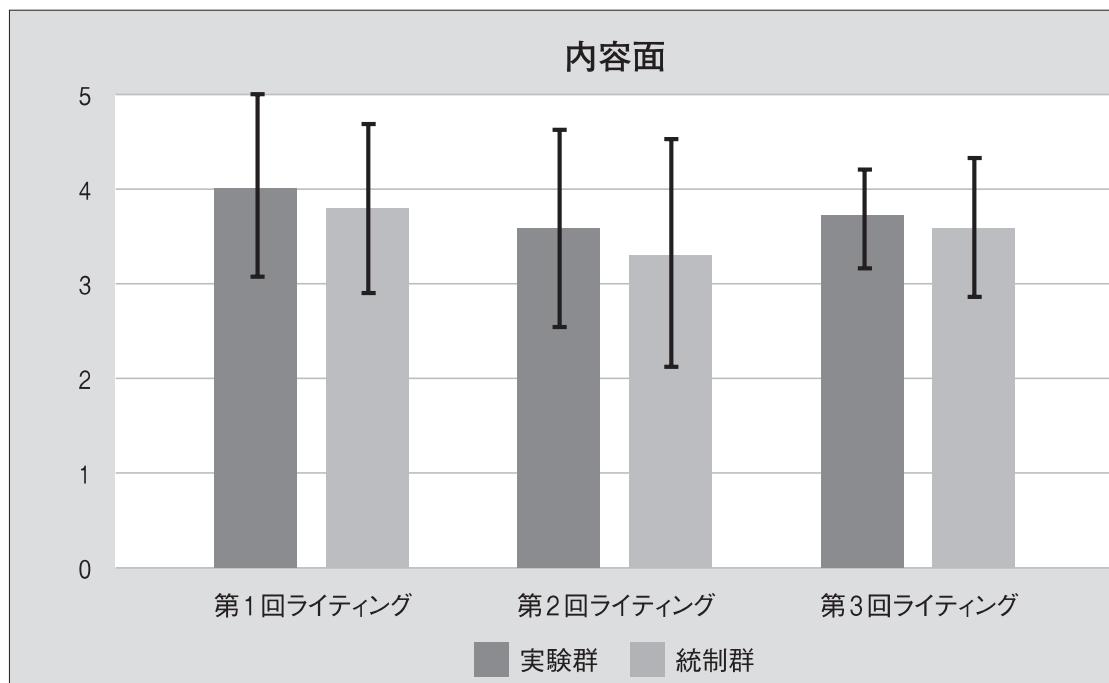
■表5: 内容面のライティングスコア

	第1回ライティング		第2回ライティング		第3回ライティング	
	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差
実験群	4.050	0.958	3.600	1.030	3.717	0.510
統制群	3.800	0.900	3.358	1.190	3.600	0.723

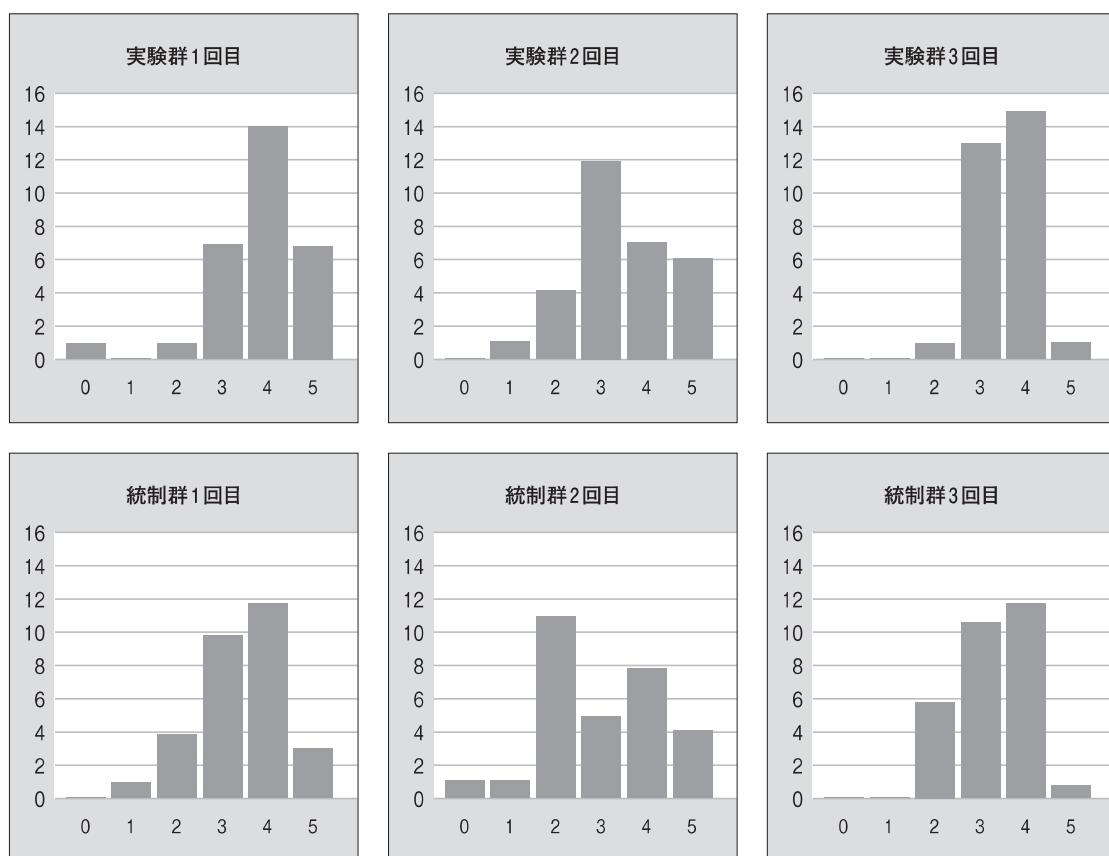
4.1 内容面のライティングスコアの結果

表5に内容面のライティングスコアを示した。図3と図4は、スコアのグラフとスコアの人数分布をそれぞれ表している。スコアを比較するために、「事前提示の有無」とライティングの「回数」の2つを要因とし、事前提示の有無は実験群と統制群の2水準(対応なし)、回数は1回から3回の3水準(対応あり)に設定して、混合計画による分散分析を行った。分散分析を行うための前提である等分散性を確認するために、各ライティングのグループ間において、Leveneの検定を行った。その結果、 $F (1, 58) = 0.041$ ($p = .841$), $F (1, 58) = 1.229$ ($p = .272$), $F (1, 58) = 3.115$ ($p = .083$) となり、それぞれのグループ間で、分散が異なるとは言えないことが確認されたので、分散分析を実施した。

表6は、内容面のスコアの分散分析の結果である。分散分析の結果、事前提示の有無と回数の交互作用は、 $F (2, 116) = 0.159$ ($p = .854$) となり有意でなかった。事前提示の有無についても、 $F (1, 58) = 1.219$ ($p = .274$) で有意な差は見られなかった。また、回数に関しては、 $F (2, 116) = 5.715$ ($p = .004$) で有意な差が見られた。そこで、ホルムの方法による多重比較を行なった結果、1回目と2回目に有意な差があり、両群とも1回目から2回目にかけてスコアが有意に下がっていた($t (58) = 3.110$, $p = .009$, d (効果量) = .430, $1 - \beta$ (検定力) = .374)。



■図3: 各回の内容面のライティングスコアのグラフ(Y軸はスコア。エラーバーは標準偏差。)



■図4: 内容面のスコアの人数分布(上部が実験群。下部が統制群。X軸はスコア。Y軸は頻度。)

■表6: 内容面のスコアの分散分析表

	SS	df	MS	F	p	η^2	1- β (検定力)
事前提示	1.850	1	1.850	1.219	.274	.020	.197
事前提示の誤差	88.009	58	1.517				
回数	6.039	2	3.019	5.715	.004	.090	.552
交互作用	0.167	2	0.084	0.159	.854	.003	.061
交互作用の誤差	0.1674	116	0.528				
全体	61.2931	179					

4.2 言語面のライティングスコアの比較

表7に、言語面のライティングスコアの結果を示す。図5と図6は、スコアのグラフとスコアの人数分布をそれぞれ表している。また、内容面と同様に、言語面のスコアも分散分析を実施するた

めに、Leveneの検定を行った。その結果、各ライティングにおけるグループ間の比較において、 $F(1, 58) = 0.116 (p = .734)$, $F(1, 58) = 2.563 (p = .115)$, $F(1, 58) = 1.087 (p = .302)$ となり、分散が異なるとは言えないことを確認した。

■表7: 言語面のライティングのスコア

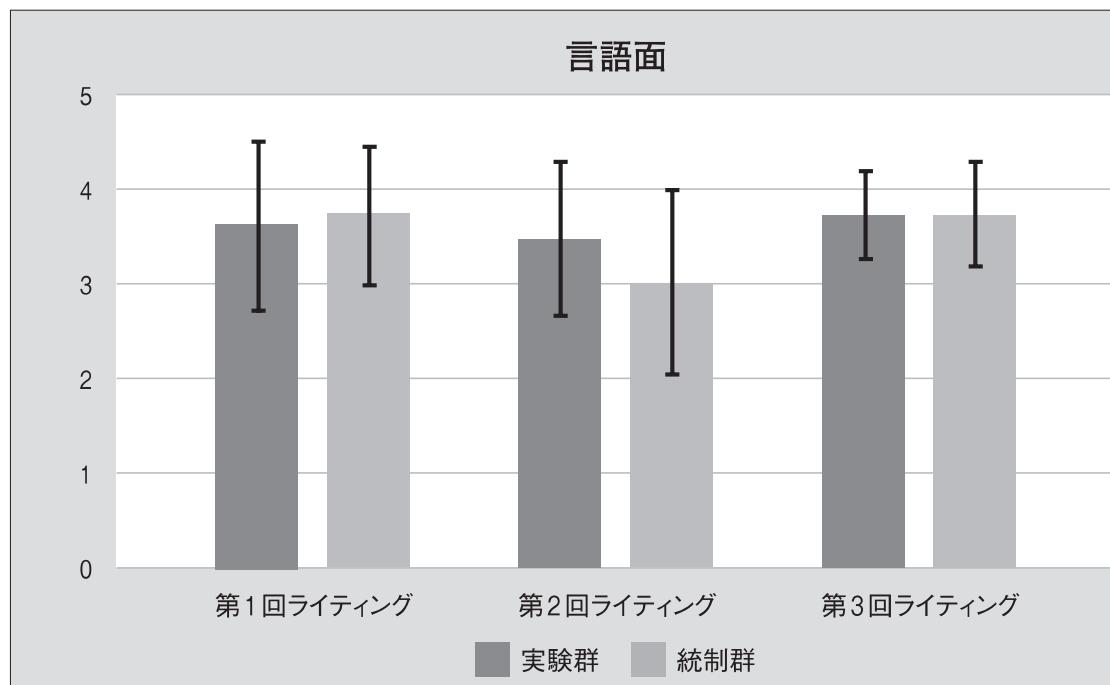
	第1回ライティング		第2回ライティング		第3回ライティング	
	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差
実験群	3.575	0.869	3.417	0.797	3.667	0.467
統制群	3.675	0.728	2.983	0.962	3.692	0.531

表8は、言語面のスコアの分散分析表である。分散分析の結果、事前提示の有無と回数の交互作用については、 $F(2, 116) = 3.054 (p = .051)$ で、有意な差は見られなかった。事前提示の有無の要因も、 $F(1, 58) = 0.519 (p = .474)$ で有意な差は見られなかった。また、回数の要因は、 $F(2, 116) = 10.093 (p < .001)$ で有意差が見られた。ホルム

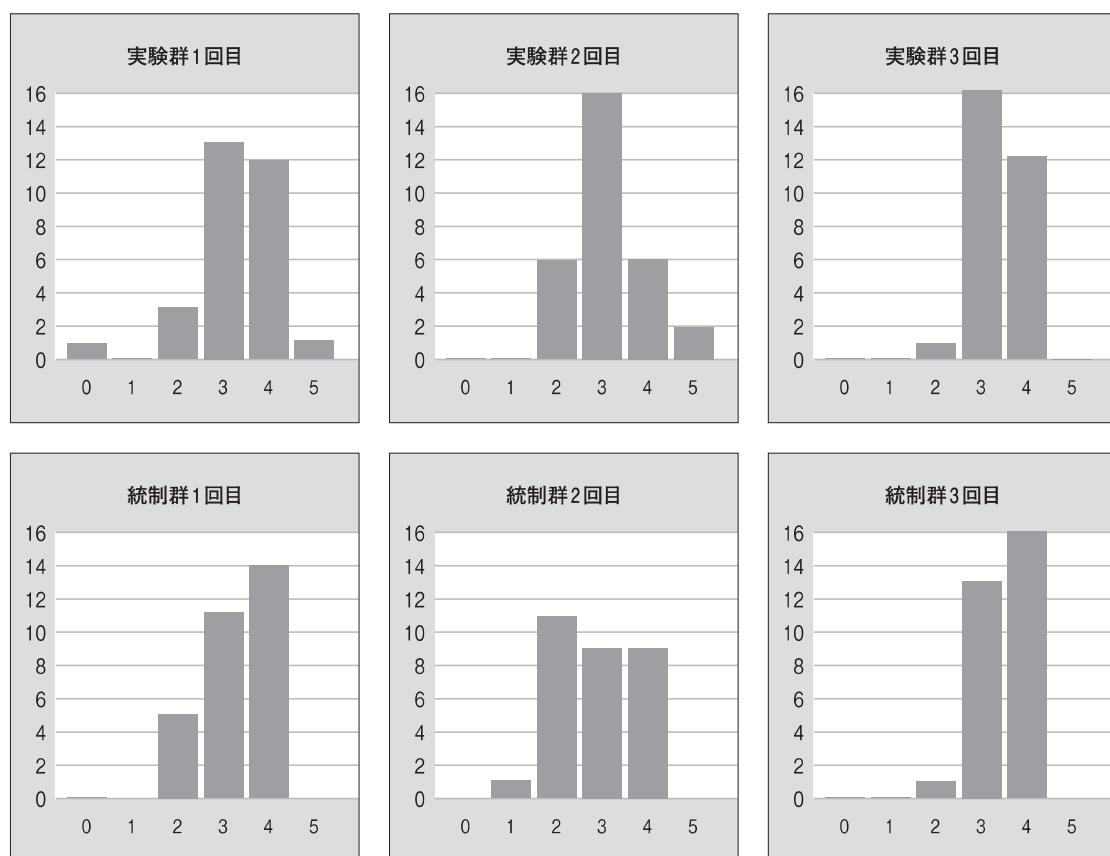
の方法で多重比較を行った結果、1回目と2回目 ($t(58) = 3.171, p = .005, d$ (効果量) = .500, 1- β (検定力) = .479), 2回目と3回目 ($t(58) = 4.428, p < .001, d$ (効果量) = .650, 1- β (検定力) = .697), で、有意な差が見られた。

■表8: 言語面のスコアの分散分析表

	SS	df	MS	F	p	η^2	1- β (検定力)
事前提示	0.475	1	0.475	0.519	.474	.009	.113
事前提示の誤差	53.097	58	0.915				
回数	8.263	2	4.131	10.093	<.001	.148	.810
交互作用	2.501	2	1.250	3.054	.051	.050	.320
交互作用の誤差	47.486	116	0.409				
全体	111.812	179					



■図5: 各回の言語面のライティングスコアの比較(Y軸はスコア。エラーバーは標準偏差。)



■図6: 言語面のスコアの人数分布 (上部が実験群。下部が統制群。X軸はスコア。Y軸は頻度。)

4.3 アンケート結果

表9にアンケートの結果を示す。前節でも述べたようにループリックに関する質問を含んだ項目の(1)～(16)を実験群に、また(6)～(16)を統制群

に回答してもらった。すべての質問に生徒は5段階で回答した。その結果、フィードバックについて聞いた項目(12)と(13)を統合して考えると実験群の方が高い数値を示していると言える。しかし、その他の項目で顕著な違いは見られなかった。

■表9: 事後アンケートの項目と記述統計

	実験群		統制群	
	平均	標準偏差	平均	標準偏差
(1) 授業中に行ったループリックの説明は理解できましたか。	4.53	0.62		
(2) その際に、例として提示されたWritingは、なぜその評価になるのか理解できましたか。	4.30	0.59		
(3) ループリックの評価の記述はわかりやすかったですか。	4.30	0.78		
(4) ループリックはWritingを書く上で、役に立ちましたか。	4.27	0.85		
(5) Writing中に、ループリックの評価基準を確認したり、意識したりしましたか。	4.33	0.65		
(6) Writing中に、内容が筋の通ったWritingにしようと意識しましたか。	4.67	0.47	4.47	0.72
(7) Writing中に、全体の構成を良くしようと努力しましたか。	4.47	0.56	4.50	0.56
(8) Writing中に、文法のミスを少なくしようと努力しましたか。	4.48	0.56	4.47	0.88
(9) Writingの中で、いろいろな語彙を使おうと努力しましたか。	4.31	0.79	4.13	0.92
(10) Writingをしている際に、良い評価を取りたいと思いましたか。	4.41	0.72	4.50	0.67
(11) 返却された答案が、その評価になる理由が理解できましたか。	4.14	0.78	4.30	0.90
(12) 返却された1回目の評価は、2回目のWritingに、役に立ちましたか。	3.97	0.76	3.77	0.84
(13) 返却された2回目の評価は、3回目のWritingに、役に立ちましたか。	4.17	0.70	3.79	0.76
(14) 計3回のWritingは英語の勉強になったと思いますか。	4.55	0.56	4.67	0.54
(15) 3回分のWritingを書いたことで、今後もっと英語の勉強をしようと思いましたか。	4.17	0.79	4.17	0.86
(16) 今回のようなWritingを繰り返し行ったら、Writingが向上すると思いますか。	4.41	0.62	4.50	0.62

5 考察

本実践研究は、ループリックの事前提示により、課題が求めている評価規準および評価基準が生徒に透明になることで、パフォーマンスも向上するという予測のもと、英語ライティング課題を実施した。しかしながら、内容面、言語面ともにループリックの事前提示の有意な効果は示されなかつた。以下に、ループリックをどのように使えば良いかという観点から、本実践で掲げた研究課題について、結果を踏まえて考察する。

5.1

RQ1: ライティング課題において、ループリックを事前に提示した場合、そうでない場合と比較し内容面においてパフォーマンスが向上するか

3回のライティングの内容面において、1回目のスコアは2回目より有意に高かったが、これは2つの解釈ができる。第1に、採点に関して、特に回数の間で評価にぶれが生じた可能性は排除できない。今回の実践では、ライティングの間の期間が3週間ほどあるため、ループリックを用いても回数により評価の判断が異なった可能性がある。実際に、2人の採点者の一致度は、1回目は、 $k = .682$ であったが、2回目、3回目は $k = .460$, $k = .467$ となり、かなりの一一致から、中程度の一一致となっている。

それが1回目のライティングが2回目よりも有意に高い結果となった1つの原因となったことはありえる。第2に、生徒の観点から見ると、1回目のフィードバックを踏まえて、さらに上の評価を目指そうと書いたゆえに、向上の過程として2回目のスコアが下がったのかもしれない。表10は、3回のライティングの内容面のスコアの推移を表している。この表を見ると、2回目でスコアが下

がったが、3回目でその2回目より上がった生徒は両群あわせて20名おり、他のケースよりも比較的人数が多い。このことは、実験群、統制群とともに良いものを書こうとして、2回目でスコアを下げてしまった可能性を間接的にではあるが支持している。今後、生徒のライティングの詳細な分析により、この過程を分析したい。

■表10: 3回のライティングの内容面のスコアの推移

	統制群	実験群
3回を通して下がらず上がった	4人	4人
2回目で下がり、3回目で2回目より上がった	11人	9人
3回を通して上がらず下がった	6人	10人
2回目で上がって、3回目で2回目より下がった	8人	5人
3回全てで変わらない	1人	2人

本実践の研究課題の1つ目である内容面のルーブリックの事前提示の効果については、実験群と統制群の比較においては、有意な差が見られず、その効果は確認できなかった。この点に関しては、以下の3点の可能性を検討する。

第1に、今回のルーブリックの記述は、本実践に参加した生徒にとって、具体性が弱く、必ずしも明確な基準となり得なかつたことが考えられる。Alter and Chappuis(2006)は、ルーブリックが満たすべき要件として、内容や組織と、明瞭性と

いう2つを挙げており、明瞭性については、「特定化しにくい言葉を使わないようにしているか」としている。ルーブリックの基準の理解について質問をしているアンケート項目の(1)～(5)への実験群の生徒の回答は、4.27～4.53の範囲にあり、全体としては基準を理解できたものと解釈できる。また、自由解答でも「ルーブリックを意識できた」「評価が分かりやすかった」と記述している生徒も7名いた。一方で、以下のようにルーブリックの曖昧さを指摘する生徒もいた。

- 評価4と評価5の違いがわかりにくかった（2名）
- 1段階ごとの基準の違いがわかりにくかった
- 具体的にどのように直せば良いのか、例などが欲しかった
- ダメな例があると良かった

このような意見があったことを踏まえると、一部の生徒には明瞭性が不十分だったために、事前提示が十分に機能しなかったのかもしれない。例えば、今回使用した内容面のルーブリックの「理由」に関する記述は、評価5が「具体的によく表現されており妥当」、評価4が「十分に表現されており妥当」、評価3が「概ね妥当」というように曖昧さは否めない。また、「どの教科が好きですか」というテーマのもとに書かれたアンカーについて

の説明では、「英語が好きな理由として、『海外に行きたいから』」というものは、海外の全てで英語が通じるわけではないから、理由としては妥当ではない」というような指導をしたが、このような説明だけでは、論理性や妥当な理由づけが何であるのかを十分に理解できなかつたかもしれない。このような点を考えると、「ルーブリックの評価が分かりやすかった」と答えた生徒においても、具体的な修正点など改善点が必ずしも明確であつ

たかはわからず、パフォーマンスに影響を与えたことは考えられる。

第2に、ルーブリックを提示されなかった統制群の生徒は、すでに今回のライティングで何が求められているのかを理解しており、ルーブリックを提示された実験群と同じ状況にあった可能性もある。アンケートの結果を見ると、「内容を筋の通ったものにしようと努力したか」、「構成をよくしようと努力したか」といった内容面に関する質問のスコアは、前者は0.2ポイントほどの違いはあるものの統制群においても4.47という比較的高い結果であり、両群で大きな違いがあるとは言えない。統制群においては、「筋の通った」「論理的な」「構成」といった言葉は、実践中には使用せず、アンケートで初めて質問した要素である。それにも関わらず、こうした項目への回答が、実験群と同様の傾向を示すことは、今回のようなエッセイライティングでは、理由づけとその一貫性が大事である、ということをルーブリックの提示がなくても意識できていたかもしれない。また、同様に、統制群に対するフィードバックでは、点数のみが提示されるだけであるが、アンケート調査では、「返却された答案が、その評価になる理由が理解できましたか」という質問に実験群よりもやや高い回答をしている。こうした点からも、ルーブリックを提示されなかったものの、評価として何が求められているのかを理解していた可能性を考えることができる。今後、学習者がどのように事前提示やフィードバックを解釈し、パフォーマンスに活用しているのかをより深く理解するために、本実践で実施したような選択肢式のアンケートではなく、インタビューや記述式のアンケートを用いて調査することが必要であろう。

第3に、そもそも事前提示だけでは、パフォーマンスに影響を与えるほどの効果を得られない可能性もある。実験群は、初回のライティングの授業において、ルーブリックと各尺度におけるアンカーを提示され、評価基準の説明を受けた。また、1つの答案について、どの評価に当てはまるかを検討してもらった。しかし、そうした事前提示だけではルーブリックや、評価基準に対する理解は十分得られず、例えば、フィードバックの際に、再度ルーブリックの記述について説明を加えたり、生徒の答案をアンカーとして説明をしたりす

ることで、パフォーマンスに影響を与えるくらいの効果が表れるのかもしれない。また、Goodrich Andrade(2001)が指摘するように、ルーブリックにより、良いライティングに関する理解が高まったとしても、ただちにスコアに反映されないのかもしれない。一方、Goodrich Andrade & Boulay(2003)は、ルーブリックの提示の効果が大きい可能性に言及しているが、2回の推敲後のスコアを分析しているため本研究と単純に比較することはできない。さらに、寺嶋、林(2006)など多くの研究がなされているように、自己評価が有効であるかもしれない。事前提示に加え、下書きの時点や、課題を完了した時点、またフィードバックを踏まえて、自己評価をすることで、ルーブリックの記述と自分が足りていないもののギャップに気づき、それが次のパフォーマンスに活かされる可能性は今後さらに検討されるべきである。相互評価に関しても同様の検証が必要である。また、ライティングに関しては、その向上にはもともと時間がかかるため、今回の実践のように2ヶ月弱での実践ではなく、長期的に実施することで、事前提示も意味があるものになってくるのかもしれない。形成的な評価という観点からも、事前提示が長期的にライティングの向上に及ぼす影響も、今後検証する必要があると言える。

5.2 RQ2: ライティング課題において、ルーブリックを事前に提示した場合と、そうでない場合を比較し言語面においてパフォーマンスが向上するか

言語面においても、2回目のスコアが1回目と3回目と比較して、有意に高い結果を示した。内容面でも述べたように、これには採点の信頼性と生徒のライティングの向上の過程という2つの可能性が考えられる。言語面の評価の信頼性に関しては、1回目は一致度が「まずまずの一致」で、3回目に関しては採点者の相関関係は、 $r = .236$ 、一致度も「わずかな一致」にとどまった。このように、2人の採点者のずれが2回目のスコアが有意に低くなったことに影響していることはあり得るだろう。また、ライティング向上の過程に関して、内容面同様に表11に、3回の言語面のスコアの推移を示した。2回目で下がり、3回目はその2回目より上がった生徒は両群で24名おり、内容面同様の傾向を示した。

■表11: 3回のライティングの言語面のスコアの推移

	統制群	実験群
3回を通して下がらず上がった	8人	6人
2回目で下がり、3回目で2回目より上がった	14人	10人
3回を通して上がらず下がった	6人	7人
2回目で上がって、3回目で2回目より下がった	0人	5人
3回全てで変わらない	2人	2人

実験群と統制群の比較には有意差は見られず、ルーブリックの提示に有意な効果は見られなかつたが、言語面に関しては、その要因として以下の4点が考えられる。

第1に、内容面と同様に、ルーブリックの記述が明瞭ではなかったのかもしれない。特に、言語面に関しては、実験群の初回のライティング時に、どのようなライティングが言語面において評価されるのかについて詳しくは説明をしなかった。また、言語面に焦点をおいたアンカーも提示しなかった。このため、例えば、言語面のルーブリックの中の「文の作り方」「語彙の豊富さ」と言った記述の具体的な内容がわからず、パフォーマンスに活かされなかつたのかもしれない。

第2に、中学1年生にとって、ライティング課題そのものが負荷の高い課題であり、言語面を意識することができなかつたのかもしれない。本実践研究に参加した生徒は、中学校1年生から英語を本格的に学習した者がほとんどであった。対象となった生徒は、入学以降、比較的早い段階から、英語の授業の中で、自己紹介や自分の好きなものの紹介といったライティングをしてきた。そのため、書くことにはある程度抵抗のない状態であったと推測できるが、自分の意見を理由と共に主張するタイプのライティングの経験は浅く、20分で自分の主張をする課題は負荷が高かつた可能性がある。このために、意識はしていたとしても、言語の形式や語彙の種類、文法などの正確さに注意を払えなかつたのかもしれない。中学1年生に対しては、パフォーマンス課題の難易度が低い自己紹介などのテーマであれば、言語知識にも注意が向けることが可能となり、言語面への事前提示効果が見られるのかもしれない。

また、第3に、ルーブリックの言語面の評価で求められていることを実践しようとしたが、その基準を満たす言語知識を持ち合わせていなかつた可能性もある。本実践は、1回目のライティングから3回目のライティングまでの期間は、6週間しかなく、言語面に注意を払ったものの評価に見合う言語知識がまだ獲得されておらず、パフォーマンスに反映できなかつたのかもしれない。例えば、辞書で単語を調べて使おうとしても、品詞の知識はまだ浅いため、その語の使い方がわからず、使用を躊躇ってしまう、というケースなども想定される。学年が上がった中学2年生や中学3年生に対して同様のライティングを課した場合、事前提示が、蓄積された言語知識や、調べた語彙や表現などを幅広く使用することを活性化させるかもしれない。

以上のように、第1~3で述べた可能性は、パフォーマンス評価を実施する際は、課題のレベル、ルーブリックの記述、学習者の習熟度の3点をバランスよく検討する必要性を示唆している。

第4に、事前提示は言語面のパフォーマンスの向上に効果を与えるほどの効果がない可能性もある。ルーブリックの機能の1つは、評価を透明にすることであるが、言語面の評価の透明性においては、ルーブリックの記述をより具体的にしたり、フィードバックも個人個人に合わせたコメントを付与したりするなどの形で実施しなければ、パフォーマンスに効果が見られないのかもしれない。特に、ライティングのフィードバックにおいては、どのようなフィードバックが効果的か、そもそも言語面のフィードバックは有効なのか、について研究結果は一致していない(田中, 2015)ため、さらに研究を重ねる必要がある。

5.3

RQ3: ライティング課題において、ループリックを使用した場合と、そうでない場合を比較し、課題に対する目標意識や、その後の学習への意欲が高くなる傾向が見られるか

3回のライティングが終了した後、それぞれのグループにアンケートを実施した。実験群と統制群の共通の項目において、フィードバックの有用性について聞いた項目(12), (13)で若干の傾向の違いが表れた。フィードバックに関しては、統制群はスコアのみの開示であるので、フィードバックとして不十分だったと感じる生徒は実験群より多くなることは予想されたことである。

しかしながら、その他については、実験群と統制群で大きな違いが見られなかった。アンケートによる調査方法は、当然ながら解釈には制限や限界があり、実際の生徒の意識を反映しているとは限らないが、事前提示に効果があれば、それぞれの項目に関しても、回答の傾向が異なることが予測される。今回は、ライティングのスコアの比較に有意な差が表れなかつたが、今回実施したアンケート項目で聞いているような内容において、少なくとも生徒の意識の点で違いがより明確に見られてくると、パフォーマンスへも影響していく可能性もある。そういう意味では、より生徒が具体的に理解し意識できるループリック作りと例を伴う説明が、ループリックを効果的に使用する上で、重要な鍵の1つとなってくると考えられる。

6

まとめ

本実践は、中学1年生のライティング活動において、ループリックの事前提示がパフォーマンスに有効かどうかを統制群とスコアを比較することで実証的に調査することを目的としたが、その効果は示されなかつた。ループリックの事前提示においては、記述の明確さ、ループリックの説明の仕方、アンカーの提示などが効果的な提示として必要であり、またそれを複数回行うことでのり評価基準を理解することができるかもしれない。また、本実践がそうであったように、事前

提示やフィードバック、自己評価や相互評価自体のそれぞれの有効性を検証することも大事であるが、Panadro & Jonsson(2013)がループリックの使用を効果的にする要因として挙げている自己評価やメタ認知活動と組み合わせることで、ライティング力が向上するかどうかを検証することも必要である。ループリックを自己評価や相互評価とともに使うことは、「学びとしての評価」という観点からも重要であると考えられる。

また、ループリックによる「効果」をどう捉えるか、という観点も非常に重要な意味を持っている。例えば、本実践でもフィードバックについて言及し、ループリックによるフィードバックの有効性を調査したが、第2言語習得研究(SLA)と第2言語ライティング研究では、フィードバックの効果をどう捉えるかが異なっている(田中, 2015)。前者は、フィードバックを受けて、そのフィードバックが新たなライティングをする際に効果を持っているかを検証するのに対し、後者は、1つのライティングのドラフトの段階でなされたフィードバックが書き直し後に効果を持っているかどうかを対象としており、ライティング全体に対する自己編集(self-editing)ストラテジーの発達過程にその焦点がある(Ferris, 2010)。この点を踏まえると、本実践では、SLAの観点からループリックの事前提示とフィードバックの効果を検証したといえるが、ループリックの性質を考えると、ライティング研究が対象にするドラフト段階での発達過程にこそループリックの提示やフィードバックが効果を持つことはありそうである。そして、その積み重ねの先にSLAで定義される「効果」が表れるのかもしれない。今後は、この「効果」をどう据えるかという点も踏まえて、ループリックをどう使うかについて検証し、実践していくことが必要だと思われる。

注

- (1) 実社会を想定して課される課題を用いて活用する力を評価することを真正の評価と呼ぶ。Wiggins(1998)は、真正を「大人が仕事場、市民生活、私生活の場で『試されている』、その文脈を模写したりシミュレートしたりすること」(p. 24)と定義している。後述のパフォーマンス評価も実際の社会での知識の活用が問

- われている課題の場合、真正な評価の対象となり、両者はオーバーラップする場面が多い。
- (2)「評価規準」は、目標に準拠した評価を行う際の目標を指す。これに対し、「評価基準」は、「規準」(目標)に従って、教師が実際に評価をおこなうときに、指標として用いるもので、例えば、「どういうものができていれば5」で、「ここまでであれば4」というように具体化された評価の指標のことである(田中, 2010)。
- (3) アメリカの教員が北西地区研究実験所(Northwest Regional Educational Laboratory)の支援のもと開発された小学校3年以上を対象にした書き方のためのルーブリック。
- (4) ライティングのフィードバックには、内容や構成についての「内容面」のフィードバック(content-based feedback)と、文法や語彙などの「形式面」のフィードバック(grammar-based feedback)の2つがあるとされている(田中, 2015)。

謝 辞

本研究への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。「研究はひっそりやるもの」という小池先生の言葉を胸に、日頃の指導と共に本研究を進めて参りました。そしてその中で、常に的確にかつ今後の指導や研究にもつながるような助言を下さりました指導助言者である小泉利恵先生には深く感謝申し上げます。ありがとうございました。また、多忙な中、本研究の採点を快く引き受けていただいた中央大学附属中学校・高等学校の園田恒子教諭および小塙英男教諭に改めて御礼申し上げます。ありがとうございました。

参考文献(*は引用文献)

- * 安藤輝次 (2014). 「ループリックの学習促進機能」『関西大学文学論集』, 64(3), 1-26.
- * Arter, J., Spandel, V., Culham, R., & Pollard, J. (1994). The impact of training students to Be self-assessors of writing · paper presented at the Annual Meeting of American Educational Research Association (New Orleans, LA, April 4-8), 2-20.
- * Arter, J., & Chappuis, J. (2006) *Creating & recognizing quality rubrics*, Educational Testing Service.
- * Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, 32, 181-201.
- * 藤森千尋・小泉利恵 (2011). 「スピーチ発表におけるALTとJTEによる全体的評価:正確さ・流暢さ・複雑さの客観的測定と生徒による相互評価と比較して」 *KATE (Kanto-koshinetsu Association of Teachers of English) Journal*, 25, 21-31.
- * Goodrich Andrade, H. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education* 4(4).
- * Goodrich Andrade, H., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research*, 97 (1), 21-34.
- * International Baccalaureate (2011). *Language B guide*. Cardiff: International Baccalaureate Organization (UK) Ltd.
- * 石井英真 (2015). 「教育目標と評価」西岡加名恵・石井英真・田中耕治 (編). 『新しい教育評価入門一人を育てる評価のためにー』(pp.77-111). 有斐閣コンパクト
- * 石井英真 (2016). 「資質・能力ベースのカリキュラムの危険性と可能性」『カリキュラム研究』25, 83-89.
- * 松井市子 (2015). 「評価rubricを活用した英語ライティング力と自己評価力の育成をねらった実践」 *EIKEN BULLETIN*, vol.27, 146-164.
- * 松下佳代 (2007). 「パフォーマンス評価—子供の思考と表現を評価するー」日本基準ブックレット No.7.
- * 松下佳代 (2012). 「パフォーマンス評価による学習の質の評価—学習評価の構図の分析にもとづいてー」『京都大学高等教育研究』18, 75-114.
- * 沖裕貴 (2014). 「大学におけるループリック評価導入の実際—公平で客観的かつ厳格な成績評価を目指してー」, 『立命館高等教育研究』14, 71-90.
- * Panadro, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144.
- * Stevens, D. D. & Levi, A. J. (2005). *Introduction to rubrics*. Sterling, VA: Stylus Publishing.
- * Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.) Upper Saddle River, NJ: Merrill/Prentice-Hall.
- * 鈴木雅之 (2011). 「ループリックの提示が学習者に及ぼす影響のメカニズムと具体的事例の効果の検討」『日本教育工学会論文誌』35(3), 131-143.
- * 田中耕治 (編) (2010). 「よくわかる教育評価」[第2版] ミネルヴァ書房
- * 田中真理 (2015). 「ライティング研究とフィードバック」大関浩美 (編・著), 名部井敏代・森博英・田中真理・原田三千代 (著)『フィードバック研究への招待—第二言語習得とフィードバックー』. (pp. 107-138). くろしお出版
- * 寺嶋浩介・林朋美 (2006). 「ループリックの構築により自己評価を促す問題解決学習の開発」『京都大学高等教育研究』12, 63-71.
- * 鶴田京子 (2018). 「ループリックを用いて段階を踏んだ、継続的なスピーキング活動の実践」 *EIKEN BULLETIN*, vol.30, 161-176.
- * Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve Student performance*. Jossey-Bass.
- * Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.), Alexandria, VA: Association for Supervision and Curriculum Development. ウィキンズ, G. & マクタイ, J. (2012). 『理解をもたらすカリキュラム設計—「逆向き設計」の理論と方法ー』(西岡加名恵訳) 日本標準
- * 山田嘉徳・森朋子・毛利美穂・岩崎千晶・田中俊也 (2015). 「学びに活用するループリックの評価に関する方法論の検討」『関西大学高等教育研究』6, 21-30.

Model Writing

テーマ：Which subject do you like?

Model Writing① 【内容面：評価5】

My favorite subject is English. First, I want to talk with foreign people in English. I want to understand foreign culture. Second, I like an English book. I often read one. It is interesting. Therefore, I want to study English.

Model Writing② 【内容面：評価4】

My favorite subject is English. First, I want to talk with foreign people in English. Second, I like an English book. I often read one. It is interesting. So, I want to study English.

Model Writing③ 【内容面：評価3】

I like English. First, I want to talk with foreign people in English. Second, I like an English book. So, I like English.

Model Writing④ 【内容面：評価2】

My favorite subject is English. First, I want to talk with foreign people in English. Second, I like foreign movies. So, I want to study English.

Model Writing⑤ 【内容面：評価1】

I like English. I study English every day. English is difficult. An English class is fun.

Model Writing⑥ 【内容面：評価 ???】

I like English. First, I like foreign countries. I want to visit many countries. Second, English is interesting. I want to speak English. So, I like English.

資料2：生徒が書いた実際の答案(上：内容・言語とも評価5。下：内容・言語ともに評価3).....

Writing Practice ③

Topic: Which city in Japan do you like? (日本のどの都市が好きですか)

I like Hokkaido. I have three reasons. The first reason is it is very rich in nature. There are many tourist spot. It is very beautiful. The second reason is the food is very delicious and there are many types of food. The third reason is there are many valuable animals in Hokkaido. You can see animals in the city. It is really different from Tokyo. So I like Hokkaido.

Writing Practice ③

Topic: Which month do you like? (1年間でどの月が好きですか)

I like May. Because I have birthday. On birthday, I can eat delicious food and receive presents.

I don't like Hot season. So I like the temperature in May. May is very refreshing season.

B 実践部門・報告Ⅱ・英語能力向上をめざす教育実践

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

研究者:東京都／青山学院 高等部 特別教諭 江下 陣(申請時:福岡県立 小倉南高等学校 教諭)
《研究助言者:和泉 伸一》

概要

本研究は学習者がスピーチを行った後に、モデルスピーチと比較することでどの程度主体的な「気づき」が起き、その後のスピーチが変容するかを実践的に調査したものである。日本人高校1年生10名の実験協力のもと、スピーキングタスク（英検準2級面接問題Q2・Q3）を行ってもらい、自身のスピーチとモデルスピーチを比較させ、気づいたことを筆記ランゲージング（言語面や内容面に対して学習者が気づいたことや考えたことを筆記という形で文字化して、理解を深めさせる手法）という形で言語化させた。その後、直後と1週間後にそれぞれスピーチを再度行ってもらった。筆記ランゲージングで表出された言語に対する「気づき」を分析し、指導介入した後のスピーチの変容を流暢性と正確性の観点で評価・分析した。結果としては、学習者はモデルスピーチとの比較で語彙的な要素に気づくことが多く、次に文法面に注意を向けていた。さらに、気づいたことと実際のスピーチを分析した結果、英検準2級のQ2でターゲットである現在進行形や、Q3のターゲットである、soやbecauseの接続詞といった文法形式にも気づいた可能性のある学習者もいたことがわかった。さらに、学習者の指導介入後のスピーチの流暢性と正確性は両方とも向上し、モデルスピーチとの比較+筆記ランゲージングの手法にスピーチの質を向上される効果がある可能性が示唆された。本研究はスピーキング指導に対して、新しい指導方法を提案している。学習者のスピーチに対して即座に誤りを訂正するなどのフィード

バックをするのではなく、一度学習者自身でモデルスピーチとの比較などを通じて自分で誤りやモデルスピーチとの違いに「気づく」機会を与え、それを筆記ランゲージングなどの方法で言語化させるという主体的・自立的な学習方法を少人数での研究ではあるものの、教育的応用の可能性が高い研究結果であると考える。

1 背景

文部科学省が令和元年度に全国の中学生を対象に行った全国学力・学習状況調査で初めて英語の「話す」が導入された。その結果として、「聞く」力を問う問題の平均正答率が68.3%、「読む」力の正答率が56.2%、「書く」力の正答率は46.4%，最後に「話す」力の正答率は30.8%と一番低いというものであった。この結果から質問に対して即興で答える問題や30秒程度のスピーチを行う力が弱いという分析がなされた（文部科学省、2019）。この結果と文部科学省の分析を踏まえると、今後は高等学校でも課題として浮彫となった英語で「話す」力の育成に重点が置かれることが考えられる。また、文部科学省が行った平成29年度英語教育実施状況調査の結果から「話す」力において、高校3年生の12.9%しか、CEFRでA2レベル（英検準2級程度）に到達しておらず、ほとんどの高校生の卒業時の「話す」力はA1レベル（英検3級程度）であると指摘されている（文部科学省、

2018)。この結果からも「話す」力に対する指導は必要であり、指導方法についても考案していく必要があると考える。

このような現状において、スピーキング指導の需要が増しているが、スピーキングの指導においては、個別での指導になることが多く、時間と労力を要することが課題となることが予想される。そこで、本研究のテーマである、学習者自身による「主体的な気づき」を促すことができ、スピーキング技能を向上させることができるのであれば、従来の個別の指導方法に加えて、新しい指導方法の提案をすることができるのではないかと考える。

2 先行研究

2.1 第二言語習得研究における「気づき」

Schmidt(1990)は、気づき仮説をとなえ、外国语の学習において気づきが必要不可欠であり、インプットは言語に対する「気づき」が起きない場合にはインテイクにならないと主張した。その後、様々な議論がされる中で、Schmidt(1994)は、“more noticing leads to more learning(p. 129)”と述べており、気づきが学びを促進すると主張した。もちろん気づいたものすべてが習得に結びつくということではないが、「気づき」は学習者の中間言語の発達において重要な役割を果たしている。さらに、どの程度意識が言語に向いているかの違いもあり、Leow(2015)では気づきの深さの違いを Depth of Processing と呼んでいる。学習者の気づきを考察する際にも、どの程度の「気づき」なのかを考える必要がある。そのような議論を踏まえても、本研究の狙いは、教えてもらうという受け身的な学びではなく、学習者の「主体的」な気づきを促す方法を提案することであり、自己主導的な「気づき」は習得に向かう過程で大切であると位置づけている。

気づきの種類は1つではなく複数あり、Izumi(2013)はそれぞれの関係を議論している。そこであげられている気づきの種類についていくつか説明する。1つ目に、Noticing a form (-meaning-function) relationship であり、これ

は Schmidt の気づき仮説で言われている気づきである。言語の形式と意味と機能がインプットの中でどのように表現されているのかに気づくということである。2つ目は、Noticing the gap between IL (Interlanguage) and TL (Target Language) である。IL とは、中間言語と呼ばれ、母語話者ではない外国语学習者が持つ言語知識であり、TL とは目標言語で、本研究の場合は英語である。この気づきは学習者の持つ言語知識と、ターゲットである言語で使用される表現との違いに気づくことである。3つ目は、Noticing holes in IL である。この気づきは、自分が表現したいことが自身の持つ現在の言語知識(中間言語)では表現できないことに気づくことである。これは直接的には言語習得に影響を与えないが、与えられるインプットから必要な言語情報を探そうとするプロセスに寄与するものであると考えられている。Izumi(2013)は、3つ目の Noticing holes in IL が、1つ目の Noticing a form (-meaning-function) と 2つ目の Noticing the gap between IL and TL の2つの「気づき」をするきっかけとなり、実際に後半2つの気づきで言語知識を増やしていくという関係性を述べている。

また、注意が向きやすいものと、向きにくいものがある。それには卓立性(Saliency)が関係している。卓立性にはいくつかの種類があるが Perceptual Saliency(知覚的な卓立性)と呼ばれるものが代表的である。この場合の卓立性を和泉(2016b)は、リスニングの際に強意が置かれる語や、リーディングでは長い語が顕著さ(卓立性)が高くなると説明している。また、文の中の語の位置も関係し、文中よりも文頭や文末の語の方が卓立性は高く、焦点が行きやすいという主張もある(VanPatten, 2007)。つまり、見た目として、どの程度その語が目立つかが Perceptual Saliency である。また、別の卓立性もある。Semantic Saliency(意味的な卓立性)とは、VanPatten(2007)のインプット処理理論(学習者がインプットをどのようにして処理するかの原則を説明したもの)に基づくもので、語の特徴からくる卓立性であり、語が発話される強意や、文中における語の位置による卓立性の高さというわけではない。インプット処理理論は、内容語

の方が文法形式より先に処理されるという The Primacy of Content Word Principle など、どの順序で学習者がインプットを処理するかについて、原則を示したものである。また、N. Ellis (2005) や Fukuta (2016) によると学習者は言語の「形式」よりも「意味」の方に注意が向くとされている。つまり、「意味」の方が卓立性が高いということである。N. Ellis (2005) が使用した例としては、Today (意味を持つ語) の方が三人称単数現在や複数形のs (文法的機能しか持たない語の部分) より卓立性が高いと具体的に述べている (Today という語は時間を表す副詞として用いられる場合には文頭や文末に位置することが多く、さらに s よりも長いことから Perceptual Saliency も高い)。

2.2 本研究の学習方法の SLA 研究からの妥当性

ここでは、本研究の一連の学習方法が SLA 研究の観点からみて妥当であることを説明し、それぞれの手順で学習者にとってどのような学習効果があるかについて触れる。便宜上、学習者のはじめのスピーチ (分析では「事前テスト」と銘打っている) を【ステージ1】とし、モデルスピーチとの比較+筆記ランゲージングを【ステージ2】、そしてその後の学習者のスピーチ (分析では「直後テスト」と銘打っている) を【ステージ3】とする。

2.2.1 【ステージ1】

学習者のスピーチの効用

学習者がまずは自分でスピーチをしてみると非常に大切であると考える。これは、Swain (1995) のアウトプット仮説でも言われていることである。Swain はアウトプットの機能の1つとして、気づきを促すと主張している (noticing function of output)。つまり、アウトプット (ライティングやスピーキング) の際に自分の言いたいことが目標言語 (本研究の場合は英語) で言えないという自身の言語能力の「穴」に気づき、その気づきが次のインプット時の気づきに貢献するということである。このアウトプットの効用は Izumi (2002) などでも確認されている。インプットにより学習者の注意を向ける方法として

Input Enhancement (注目してほしい言語情報を太字や色を変えるなどする方法) があるが、それよりもアウトプットさせる方が効果的であることも述べられている (Izumi, 2002)。また、「穴」だけでなく、インプットと自分のアウトプットとの違い (ギャップ) に気づくことにも言語習得に貢献すると言われている。本研究においてもはじめに学習者自身が上手にスピーチできないかもしれないが、とりあえずスピーチをさせることは非常に重要な役割を担っている。さらに、はじめのスピーチが中途半端にできずに終わってしまうことも次の学習意欲に繋がる可能性もある (Izumi, 2003)。

2.2.2 【ステージ2】

モデルスピーチとの比較+筆記ランゲージングの効用

自身のスピーチとモデルスピーチを比較し、学習者自身で主体的に「気づく」ことの効果は大きいにあることが考えらえる。Ellis, Basturkmen, and Loewen (2001) の研究では言語形式について気づいたもののうち、学習者自身で気づいたものを学習者は次のスピーチにより多く取り入れると報告している。

ランゲージングの効果については Swain (2006) が詳しく説明している。学習効果を述べる前に、ランゲージングが何であるかを述べておきたい。Swain によるとランゲージングとは、"the process of making meaning and shaping knowledge and experience through language" (Swain, 2006, p. 89) と定義している。つまり、学習者が言語に対して思考したことを言語化し外在化させることで、知識などを体系化させる過程ということである。学習者は思考したものを自分の言葉で言語化し、その思考したものを振り返ることができ、思考したことを深められると言われている。Swain (2006) 以前の研究では、ランゲージングは Metatalk (Storch, 2008), Verbalization (Swain, 2000), Collaborative Dialog (Swain & Lapkin, 1998) と研究によって異なる呼び方で呼ばれていたが、これらの研究でも自分の思考を言語化することは認知的な発達を促すと主張されている。また、SLA の研究でしばしば用いられる Think-aloud Protocol (思考発話) や Language

Related Episodes とランゲージングの過程は類似しているものの、目的が異なる。思考発話などは実験方法という位置づけで、そのプロセスのもつ効果は対象としていない（もちろん影響を考慮に入れて考察している場合も多くある）。一方でランゲージングはそれ自体を学習として捉えている。

ランゲージングの研究はこれまで口頭によるものが多くあったが、近年は筆記によるものも多くみられる（Ishikawa & Suzuki, 2016; Ishikawa, 2018）。本研究では筆記ランゲージングを用いる。筆記ランゲージングとは、主にタスク中やタスク後などに、言語面などについて思考したことをメモとして書くという、口頭でのランゲージングの筆記版である。例えば、英作文を書いている時に、表現したいけど表現方法が分からぬ表現を筆記したり、書いた後に、モデル文と比較して気づいたことを筆記したりして、認知的な発達を促すものである。筆記ランゲージングの効果は、口頭と比べて自分のペースで時間をかけて行うことができるという点にある。時間をかけ、じっくりと自分の思考したことを考え、それを振り返ることができる。

このモデルスピーチとの比較+筆記ランゲージングは過去の研究でもライティングのモードで調査されている。Hanaoka(2007)の研究では、テーマについてライティングした後に、モデル文との比較と比較の際に気づいたことをNote-takingさせ、どのようなことに注意が向くかを調査している。多くの学習者は語彙に注意を向けており、気づいたことを次のライティングに取り入れていた。なお、Hanaokaの研究ではNote-taking自体の効果を調査の目的とはしていなかった。筆記ランゲージングの効果を直接的に調査した研究もある。Ishikawa(2018)の研究では仮定法過去を含むディクトグロースタスク（読み上げられる音声を聞きながら、メモを取り、ペアなどで元の文に近づけるように再現する活動）の後に、修正された文との比較で思考したことを筆記ランゲージングし、仮定法過去の定着度合いを検証した。結果としては筆記ランゲージングを行った群において、長期的な効果が見られた。この結果についてIshikawaは筆記ランゲージングを行ったことで学習者が文をより深く処理することが

でき、その結果、長期的な記憶の保持、あるいは定着に結びついたのではないかと分析している。モデル文との比較はフィードバックと比較して、優れている部分もあるという考え方もある。Kang (2020)は、エラーに対してのフィードバックはエラーの修正を行うという部分では効果的ではあるが、モデル文との比較は内容・語彙・形式面に対して学習者に様々な提案を行うことができると主張している。

これらの先行研究から、モデルスピーチとの比較+筆記ランゲージングという学習方法を本研究の中核として捉え、この2つを組み合わせることでどのようなことに気づき、スピーチを変容させるかについて研究を進めていった。

2.2.3【ステージ3】

モデルスピーチとの比較+筆記ランゲージング後の学習者のスピーチ

モデルスピーチとの比較の後にもう一度スピーチを行うことにも意義があると考えている。この考えはフィードバックの研究からの知見であるが、フィードバック後の修正(revision)は非常に大切で、フィードバックが明示的か非明示的かの問題よりも重要な条件であると言われている。例えば、Chandler(2003)やShintani, Ellis, and Suzuki(2014)の研究ではスピーチではなく、筆記であったが、フィードバック後に学習者が書き直した方が書き直さないより効果的であると実証した。ゆえに、本研究においてもモデルスピーチの直後にもう一度スピーチを行うようにデザインした。

2.3 発話の流暢性

流暢性の研究はこれまであまりされてこなかつたが、流暢性は学習者の言語運用能力の高さを示す指標として重要な要因の1つである。流暢性はTavakoli, Nakatsuhara, and Hunter (2020)の定義では“producing language at an adequate speed with more ease and less hesitation” (p. 1)とされており、やはり流暢性は言語運用能力と高い正の相関があると主張されている。Skehan (2003) や Tavakoli and Skehan (2005) は流暢性の指標を3つの観点で考えている。1つ

目が speed fluency, 2つ目が breakdown (dys) fluency, 3つ目が repair fluency である。Speed fluency はつまり発話の速さであり、一定の時間内にどれだけの発話量があるかによって流暢さを評価するものである。これはどれだけ言語知識が自動化でき、考えたことを素早く発話できるかに関係している (Kahng, 2014)。次の breakdown (dys) fluency は発話の中でどれだけ不自然な位置での停止があるかを評価している指標であり、停止が多いと流暢性は低いということになる。最後に、repair fluency は、一度発話をしたことをどれだけ言い直したかを評価しており、言い直しが多いと流暢性は落ちることになる。本研究において、分析に扱う指標として、主に speed fluency に着目するが、個別の事例を見していく際には repair fluency の観点も見ていき、流暢性を総合的に評価し、考察していく。

2.4 発話の正確性

正確性も流暢性と合わせて学習者の言語運用の熟達度を測る指標になる(複雑性も指標であるが、本研究では考慮に入れていない)。Housen and Kuiken (2009) の定義によると、正確性はエラーのない表現を産出できる能力であり、当然エラーの少ない発話の方が熟達した第二言語ユーザーということである。また、正確性(と複雑性)は学習者のその段階での言語知識(宣言的・手続き的知識)の度合いを表している。多くの研究では正確性を error per t-unit や error free t-unit などで産出しており、t-unit という語の塊の中のエラーの割合やエラーを含まない割合で評価している。

2.5 本研究の意義

本研究は、スピーキング指導の1つの形を提案することを目的としている。英検の二次面接を含め、他の検定試験や学内で行われるスピーキングのパフォーマンステストの準備には多くの時間を要することが考えられる。スピーキングの特性上、集団指導よりも個別指導を中心に行われ、インタビュー形式で指導が行われることが多いのではないだろうか。本研究で実践している、【ス

テージ1】まずはスピーチを学習者自身がしてみる(アウトプット)、【ステージ2】自身のスピーチとモデルスピーチの比較+気づいたことを筆記ランゲージング、【ステージ3】再度スピーチを行う、の学習手順においては指導者の介入も少なく、学習者が個人で行うこともできる。この学習方法で学習者が自発的に言語面に関してなにからかの「気づき」を起こし、自身の言語知識に取り入れる(インテイクする)ことができたら、スピーキング指導の1つの形になることが期待できる。学習者は、自身のアウトプット(はじめのスピーチ)でなんとか英語で表現しようとし、その後のモデルスピーチのインプットをリソースとして使い、気づいたことを筆記ランゲージングの形で言語化する。このような主体的な学習の効用について、和泉 (2016a) でも、「自律的学習者に欠かせない要素であり、そういった学習者が、教室外でも起こる言語習得を豊かなものとしていけるのである(p.152)。」と述べており、学習者の英語を学ぶ際の態度にも影響を与えることが考えられる。

Hanaoka (2007) の研究はライティングにおけるモデル文によるフィードバックの有効性を提案した大変有意義な研究であり、その後、ライティングとモデル文の関係の研究は多くされている (Hanaoka & Izumi, 2012他)。しかしながらスピーキングのモードでの研究はまだ研究例は少なく、今後の第二言語習得研究や英語教育研究において、新たな発見や提案ができる可能性がある。本研究は調査人数も少なく、統制群も設けていない。より規模の大きい研究の予備調査という位置づけで本研究を行い、その中で得られた結果を指導・学習方法として提案したい。

2.6 リサーチクエスチョン

RQ1 自身のスピーチとモデルスピーチの比較により、学習者は言語やタスクのどのような側面に気づくか。

RQ2 スピーチによるアウトプットとその後のスピーチとモデルスピーチとの比較+筆記ランゲージングにより、学習者のスピーチの流暢性はどのように変容し、長期的に効果を継続できるか。

RQ3

スピーチによるアウトプットとその後のスピーチとモデルスピーチとの比較+筆記ランゲージングにより、学習者のスピーチの正確性はどのように変容し、長期的に効果を継続できるか。

3**調査****3.1 対象者**

本研究の対象となる学習者は、筆者の勤務する県立学校の高校1年生10名である。すべての学習者が実用英語技能検定準2級の一次試験に合格しており、面接練習の一環としてこの実践に協力してもらつた。実験の参加を承諾してくれた

10名の学習者のうち、男性は6名で、女性は4名であった。すべての学習者は日本の公立中学校で3年間、学校の授業で3年間英語を学んでおり、高校に入学してからは、コミュニケーション英語Iを4単位、英語表現Iを2単位と週に6時間英語の授業を受けている。

本研究に協力してくれた学習者の英検CSE4技能総合スコアの平均(表1)は1824.8($SD = 60.94$)であり、最もスコアの高い学習者のスコアは1920で、最も低い学習者は1752であった。英検CSEスコアとCEFRの相関図によると、1700-1950がA2レベルであるので、すべての学習者の英語熟達度はA2程度であることが分かる。英検準2級がA2レベルであると考えれば、多くの学習者は英検準2級に合格できる程度の英語熟達度を持っていたということである。

■表1：学習者の英語熟達度(英検準2級受験時スコア平均)

技能	n	Mean	SD	Median	CEFR
リーディング	10	476.2	25.6	474.5	A2
リスニング	10	439.6	23.01	441	A2
ライティング	10	488.9	40.82	480.5	A2
スピーキング	10	420.1	18.44	419	A2
4技能総合	10	1824.8	60.94	1805.5	A2

3.2 マテリアル

本研究は、2018年第2回の実用英語技能検定試験準2級の面接の問題を許可を得て使用した。分析の対象としたクエスチョンはQ2とQ3の問題である。Q2の問題は絵に登場する5名の人物の行動を描写する課題であり(Appendix A参照)，Q3の問題は絵に登場する人物が直面している課題や困難を描写する課題(Appendix B参照)である。Q2については、タスクを達成するには、現在進行形が適切に使用できるかが鍵となるタスクであり、X be V-ing Yというような構造が多く見られる。Q3については、行動と問題の因果関係をsoやbecauseを用いて説明することが求められる。

3.3 手順**3.3.1【ステージ1】****自身のスピーチによるアウトプット
(事前テスト)**

事前テストでは、Q2とQ3の問題を面接官の指示のもとに取り組んだ。Q2とQ3のタスクは別々ではなく、通して行った。学習者の発話は学習者の許可を得てICレコーダーで録音し、学習者自身にも後で発話を聞いてもらうと事前に伝えた。

3.3.2【ステージ2】**モデルスピーチとの比較+筆記ランゲージング**
2つのタスク終了後、学習者はまず自身の発話

を聞き、その後、モデルとなるスピーチを聞いた。このモデルスピーチは事前にネイティブスピーカーに協力してもらい録音したものである（Appendix B 参照）。学習者にはモデルスピーチは巻き戻して複数回聞いたり、止めたりしても構わないと伝えた。聞きながら自身の発話とモデルスピーチを比較し、気づいたことを自由に記述するように指示した。学習者から十分にモデルスピーチを聞いたと申し出があった時点での段階は終了とした。

3.3.3 【ステージ3】

指導介入の学習者のスピーチ(直後テスト)

モデルスピーチとの比較が終了した後、学習者は再度 Q2 と Q3 に取り組んだ。この際にも IC レコーダーで録音した。この録音は学習者には聞かせていない。

3.3.4 【遅延テスト】

直後テストから 1 週間後に同様に学習者は Q2 と Q3 に取り組んだ。この際にも IC レコーダーで録音した。すべての実験過程が終了した後にこの 2 つのタスクに関するアドバイスをした。

3.4 分析

3.4.1 筆記ランゲージングにおける「気づき」の分析

学習者が自身の発話とモデルスピーチを比較して気づいたことを筆記ランゲージングしたものを「語彙・文法・発音・方略・その他」の 5 点で分類した。先行研究 (Hanaoka, 2007) においては、「語彙、文法、内容、その他」のように分けられていたが、本研究においては、ライティングではなく、スピーキングであるため、発音という分類も加えた。さらに、学習者は英検の面接試験を意識し、「考えてから話すようにする」などといったテストを受ける上での方略のようなものにも注目し、記述していたため、「方略」という分類を加えた。すべての学習者の記述を一覧にし、それぞれの筆記ランゲージングを上記で説明した 5 つに分類し、筆記ランゲージングの分類ごとの頻度を産出した。その際、Q2 と Q3 のタスクにおいての特性を見るために、それぞれ別々で頻度の産出を行った。

3.4.2 流暢性の分析

流暢性の分析については ELAN を使用し、音声を文字化した。このソフトを使用することで発話のはじめと終わりを確認することができ、発話に要した時間を算出することができた。学習者の発話と算出した時間を利用し、各学習者の各タスクにおける流暢性の評価を行った。流暢性の指標としては、先行研究では、多くの指標が使用されているがその中でも本研究においては syllables per second の指標を用いた。Q2 のタスクにおいては、5 人の人物の行動を表すタスクのため、発話の時間を 2 通り産出した。1 つは、1 人 1 人の人物の行動を説明するのに要した時間、つまり前の人物の行動を説明した後に、次の人物の説明を始めるまでの沈黙は含めないもの（以後個別流暢性）と、2 つ目は、タスク全体、つまり最大 5 人の人物の行動を説明するのに要した合計の時間（以後、全体流暢性）を産出した。

3.4.3 正確性の分析

正確性の分析については学習者の発話をもとに、Q2 のタスクでは、評価の観点を大きく文法的正確性と語彙的正確性と分類し、評価した。さらに、文法的正確性についてはこのタスクの狙いとしている文法事項が現在進行形の正確な使用であると判断したため、現在進行形の正確さのみを評価の対象とした。Be 動詞と V-ing の使用の正確な利用につき、それぞれ 1 点ずつ与えた。語彙的正確性については X be V-ing Y という構文のうち、主語に当たる X と目的語に当たる Y を適切な語を使って説明できているかによって、それぞれの語の正確な利用につき、1 点ずつ与えた。Q2 は 5 文あるので、10 点満点（文法 5 点・語彙 5 点）で評価した。

Q3 のタスクにおいては、評価の観点は文法的正確性と語彙的正確性に加え、接続詞を正確に用いて、状況の因果関係を説明できているかの 3 点で評価を行った。まず、文法的な正確性の説明をする。「犬が汚れている」という状況を説明する “The dog is dirty” の節においては、SVC という語順が守られているかの 1 点で評価した。次に「女の子は犬を洗おうと思っている」という状況を説明する “The girl is going to(will / wants to) wash the dog” の節においては、SVO の語順

が守られているかの1点と、be going toやwillの未来を表す表現かwants toのように不定詞を用いて女の子の意志を表現できているかの1点、合計2点で評価した。語彙的な正確性においては、「The girl is going to wash the dog」の節においては動詞washを、「The dog is dirty」の節においては形容詞のdirtyを使用できているかそれぞれ1点ずつで評価した。最後に、因果関係を表す接続詞soやbecauseが適切に使用できているかの1点で評価した。つまり、Q3のタスクにおいては文法的な正確性が3点、語彙的な正確性が2点、接続詞の適切な使用が1点の合計6点で評価を行った。

3.4.4 統計分析

データの分析においてはいくつかの統計分析を行った。まず、筆記ランゲージングで分類した語彙・文法・発音・方略・その他の観点の頻度については χ^2 検定を実行し、頻度の差を分析した。流暢性と正確性については、人数が10名と少なかったため、ノンパラメトリック検定であるフリードマン検定を行い、事前テスト、直後テスト、そして遅延テストでの値の差を分析した。フリードマン検定後にそれぞれのテスト同士の多重比較も行った。多重比較の補正(False Discovery Rateの値を使用)をした上でウィルコクソンの符号付き順位和検定を実行した。

この研究は調査の対象とした人数も少なかつたため、統計分析で出た結果に加え、個別の学習者がどのようにスピーチを変容させていったかも考察の一部として見ることとした。そのため、考察として述べた点が個別の事象の領域を超えないこともあることは述べておきたい。

4 結果

4.1 RQ1 モデルスピーチとの比較の際の筆記ランゲージングから見られる気づき

筆記ランゲージングを実施した際の記述を「語彙・文法・発音・方略・その他」の5つに分類した。Q2, Q3におけるそれぞれの記述の頻度を表2に示している。全体として、「語彙→文法→方略→発音→その他」の順で多かった。Q2においては、特に語彙に関する記述は χ^2 検定の残差分析の結果、他より有意に多いことがわかる($Z = 9.66, p < .001$)。また、方略と発音は有意に少ないことがわかった(方略: $Z = -1.91, p = .05$, 発音: $Z = -3.11, p < .001$)。Q3においては、Q2と同様に語彙に対しての記述が有意に多いことがわかった($Z = 3.21, p = .001$)。また、発音については有意に少なく($Z = -1.97, p = .049$)、概ねQ2の結果と類似していた。1点異なる点としては、Q3においては、文法に関する記述も全体の中で有意に多かった($Z = 2.17, p = .002$)。

RQ1の分析結果をまとめると、Q2, Q3の両方において、全体の中で語彙が有意に多く、Q2においては方略と発音が、Q3においては発音についての記述が有意に少なかった。また、Q3においては文法についての記述も有意に多かった。これらの点について、考察の章ではこのような結果になった理由を考察していく。

4.2 RQ2 スピーチの流暢性の分析

4.2.1 Q2における流暢性

Q2の流暢性については、タスク全体の流暢性(全体流暢性)と1人の人物ごとの流暢性(個別

■表2: 各タスクにおける筆記ランゲージングでの「気づき」の分類と頻度

言語的指標	語彙	文法	発音	方略	その他	合計
Q2 頻度	34	7	2	5	1	49
有意差 p 値	.000**	.264	.002*	.005*	.000**	
Q3 頻度	12	10	2	3	2	29
有意差 p 値	.001**	.029*	.049*	.146	.049*	

*: $p < .05$; **: $p < .001$

流暢性)にわけて両方とも分析した(記述統計量については表3参照)。全体流暢性において、事前テストよりも直後テスト($Z = 2.19$, $p = .042$, $r = .49$)と遅延テスト($Z = 2.601$, $p = .028$, $r = .582$)で流暢性の値は有意に高く、直後テストから遅延テストにかけても有意ではないが向上している($Z = 1.58$, $p = .025$, $r = .354$)。個別流暢性においても同様で、事前テストと直後テスト・遅延テスト間での差は有意であった(対直後テスト: $Z = 2.46$, $p = .013$, $r = .551$; 対遅延テスト $Z = 2.46$, $p = .025$, $r = .551$)。

これらの結果をまとめると、モデルスピーチとの比較+筆記ランゲージングを通じて、事前テストから直後テスト・遅延テストで向上が見られ、直後テストで向上した値は維持され(むしろ若干向上している)、遅延テストでも事前テストより高い値であった。

4.2.2 Q3における流暢性

Q3のタスクにおいては、各テストの流暢性の値の平均は事前テストよりも直後テスト($Z = 1.72$, $p = .112$, $r = .386$)と遅延テスト($Z = 1.30$, $p = .169$, $r = .291$)で高く、直後テストから遅延テストの間では下降している($Z = 2.20$, $p = .085$, $r = .49$)。3つのテスト間の多重比較において有意な差は見られなかったが、事前テストと直後テスト・遅延テストとの間の効果量は中程度であり(対直後テスト: $r = .386$; 対遅延テスト: $r = .291$)、統計的にはある程度の向上はしていることがわかる。ただし、ここで1つ注意すべき点としては、ある学習者は事前テストで非常に短い文“*The girl will wash the dog.*”のみを素早く発話し、その際の値がどの学習者のどのテストで得られた値より高かった(2.54 syllables per second)ことである。そのため、事前テストの流暢性の値が高くなった可能性も考えられる。

■表3: 各テストでの流暢性の記述統計量

テスト名	n	Mean	SD	Median
Q2 全体流暢性				
事前テスト	10	0.72	0.23	0.72
直後テスト	10	0.99	0.15	0.99
遅延テスト	10	1.23	0.30	1.25
Q2 個別流暢性				
事前テスト	10	1.04	0.32	0.92
直後テスト	10	1.40	0.29	1.38
遅延テスト	10	1.64	0.47	1.52
Q3 流暢性				
事前テスト	10	0.85	0.64	0.63
直後テスト	10	1.37	0.42	1.28
遅延テスト	10	1.19	0.56	1.21

4.3 RQ 3 スピーチの正確性の分析

4.3.1 Q 2における正確性

まず、語彙的正確性についての結果を述べる。記述統計量(表4)からもわかるように、事前テストから直後テストにかけて向上し、その後の遅延テストでは若干の下降をしている。統計分析の結果を見ると、事前テスト・直後テスト・遅延テスト間における値の差(テスト間で語彙的正確性が向上したかどうか)を分析するフリードマン検

定($\chi^2(2) = 7.03$, $p = .029$)とその多重比較(事前テストと直後テスト・事前テストと遅延テスト・直後テストと遅延テストというペア間の比較)の結果、テスト間に有意な差は見られなかった($p > .05$)。しかし、効果量を見ると事前テストと直後テスト・遅延テストにおいては中から大程度の効果量があり(対直後テスト $r = .455$; 対遅延テスト $r = .478$)、有意な差ではないが、統計的に見ても向上していることがわかる。

文法的正確性においては、記述統計量を見ると、

事前テストから直後テストで値は向上し、遅延テストにかけてさらに若干向上している。統計分析の結果は語彙的正確性と同様の結果で、フリードマン検定の結果、3つのテスト間の文法的正確性の差は見られなかった($\chi^2(2) = .36, p = .367$)。効果量を見ると、文法的正確性においては、事前テストと直後テスト・遅延テスト間で中程度の効果量が見られた(対直後テスト $r = .306$; 対遅延テスト $r = .317$)。

Q2における結果をまとめると、両方の正確性において、事前テストよりも直後テスト・遅延テストにおいて向上が見られたが有意な差ではなかったものの、効果量は中から大程度あり、統計的に向上していることがわかった。直後テストと遅延テストにおいては若干の差はあるが、ほとんど変わらない数値であった。

4.3.2 Q3における正確性

Q3のタスクの正確性は文法的正確性と語彙的正確性、接続詞の適切な使用という側面で分析した。結果としては、記述統計からもわかるように、事前テストから直後テスト間で有意に向上し($Z = 2.85, p = .013, r = .637$)、直後テストから遅延テストにかけては若干の下降をしている($Z = 1.25, p = .206, r = .283$)。事前テストから直後テストにかけての向上の効果量は $r = .637$ と大程度であり、事前テストと遅延テストでの向上も $r = .456$ 中から大程度である。この点から、事前テストから直後テスト・遅延テストにかけて、Q3においても発話の正確性は統計的にも向上していることがわかる。

■表4: 各テストでの正確性の記述統計量

テスト名	n	Mean	SD	Median
Q2 語彙的正確性				
事前テスト	10	7.3	1.25	7
直後テスト	10	8.9	0.99	9
遅延テスト	10	8.7	1.66	9
Q2 文法的正確性				
事前テスト	10	7.6	2.46	8.5
直後テスト	10	8.7	1.95	10
遅延テスト	10	8.9	0.99	9
Q3 正確性				
事前テスト	10	3.6	0.97	3.5
直後テスト	10	5.4	0.84	6
遅延テスト	10	4.9	1.20	5

5 考察

5.1

RQ1(モデルスピーチとの比較の際の筆記ランゲージングに見られる気づき)について

筆記ランゲージングで学習者が記載したものと学習者が気づいたものとして分析したが、Q2, Q3の両方で、「語彙→文法→方略→発音→その他」の順で多かった。この結果は Hanaoka (2007) の研究の結果と類似している。Hanaoka (2007) に

おいても、語彙に関する気づきが多く起きており、学習者はモデル文から言語の語彙的な側面に多く注目し、気づくという結果であった。本研究では、スピーキング後にモデルスピーチをリスニングした上での気づきを調べたものであったが、Hanaoka (2007) の研究(ライティングした後に書かれてあるモデル文のリーディング)とこの側面においては類似した点があったことは興味深い。モデル文はモードがライティングであれ、スピーキングであれ語彙に注意がいく可能性が高いことが考えられる。また、方略と発音が有意に少なく、どのようにしてタスクを遂行するかについて

や、各語の音声的な情報についてはあまり注意が向かないと考えられる。

文法に関する気づきもQ2では全体49件のうち7件であり、他の項目との有意差は見られなかつたが、学習者がどの文法事項にどのような気づきを起こしたのか考察していく。7件のうち、2件が現在進行形についての気づきであった。個別の事例を見ていくと、ある学習者のスピーチで、“The old man have a cup of coffee.”と発話し(現在進行形が使用できていない)、モデルスピーチとの比較の中で、学習者の記述に「現在進行形を使って説明」とメタ言語的な理解を伴った気づきを起こしており、直後テストにおいては、5文すべてでbe+V-ingという現在進行形の形を適切に使用することができていた。また、「語彙の気づき」と分類したものにも「現在進行形」という文法名を使用するようなメタ言語的な気づきではなかったが、現在進行形に多少の注意が向かれている可能性があるものが多くあった。例えば、「コーヒーを注ぐ」や「猫に餌をやる」、「服を着る」という表現のうち、学習者の多くは動詞のpourやfeed, put onがはじめのスピーチではわからなかつたが、モデルスピーチとの比較で「feeding→餌をやる」や「putting on→服を着る」というように現在進行形の形で動詞を表記して記述している。これらの表現を学習者が個別の事象としてとらえているのか、あるいはもっと一般化した文法事象としてとらえているかは定かではない。説明はできないが何かルールがあると感じているのかもしれないが、この記述からはわからない。

英検準2級のQ2は現在進行形が適切に使用できているかがポイントであるが、これらの結果から高校1年生の学習者(CEFR A2レベル)で、中学校段階で現在進行形が既習事項であれば、モデルスピーチを聞かせることで現在進行形に関して多少の気づきは得られる場合もある。指導する際に、はじめに自身でスピーチをさせた後に、モデルスピーチとの比較をさせ、自分で気づく機会を与えた上で、気づかなかつた場合にInput Enhancementなどの方法で現在進行形に学習者の注意が向くように促したり、明示的に教えたりすることを提案できる。

Q3の分析において、語彙に対しての気づきの

他に文法に関する記述も全体の中で有意に多くあったという点は注目に値する。この結果は、becauseやsoなどの接続詞に関する記述を文法に分類したことが理由と考えられる。その証左として、10個の記述のうち8個が接続詞に関するものであった。Q3についてはbecauseやsoの因果関係を表す接続詞について多くの気づきが見られ、これらの語はこのタスクにおいて、意味的(Semantic)にも知覚的(Perceptual)にも卓立性が高いことが考えられる。Q3の状況を描写する際には因果関係を表す語として、接続詞は必要なので、伝達する情報が多かったことが意味的な卓立性の高さに寄与した。また、ネイティブスピーカーによるモデルスピーチでは、soやbecauseは他の語よりも強意が置かれて読んでいたことから知覚的な卓立性も高かったことが考えられる。学習者の筆記ランゲージングの記述には「becauseやsoで文を繋ぐ」のように接続詞の機能にも気づいた可能性が示唆されるようなものや、「…wash her dog because it is dirty」のように接続詞の機能について気づいたかどうかはわからないが、少なからずbecauseが現れたことには気づいているようなものもあった。この接続詞に対しての気づきを学習者はインテイクできており、事前テストでは10名中7名が接続詞を使用できていなかつたところ、直後テストと遅延テストにおいては全員がどちらかの接続詞を使用できていた。この結果から、このタスクにおいて、接続詞は意味機能が強く(soは「だから」, becauseは「なぜなら」という意味を持ち、文と文を繋げる文法機能とともに意味機能を持つ)、卓立性が高いのか気づくことができ、自身のスピーチに取り入れることができたことが理由として考えらえる。

RQ1に対しての考察を以下にまとめる。(1)学習者はモデルスピーチとの比較によって、語彙に最も注意が行き、その次に文法に注意が行く傾向がある。この結果はHanaoka(2007)などの先行研究で得られた情報と一致する。(2)現在進行形という文法事項にも学習者によって気づきの度合いや深さの差はあるが、おおむねモデルスピーチから気づき、インテイクすることができる。(3)soやbecauseなどの接続詞に対しても学習者は気づくことができたが、これはsoやbecauseが

文法事項の中でも卓立性が高いからではないかということが考えらえる。

5.2 RQ2(スピーチの流暢性の変化)について

RQ2に対して、Q2においては全体的流暢性・個別流暢性の両方で事前テスト、直後テスト、遅延テストと流暢性は向上していき、事前テストと直後テスト間ではその差は有意であったという結果であった。つまり、Q2においては、事前テストと直後テストの間で行った、自身の発話とモデルスピーチとの比較、比較の際に気づいたことの筆記ランゲージングにより流暢性が向上した可能性がある。ここで分析の対象としたのは流暢性の中でも speed fluency(一定の時間に発話できる量)だが、個別の事例に対して repairについても考察していく。Repair、つまり言い直しの回数と長さを見てみると、事前テストにおいては全学習者の全発話の中に20回(56 syllables)あったのが、直後テストでは16回(34 syllables)で、遅延テストでは16回(30 syllables)であった。多少はあるが、事前テストから直後テスト・遅延テストにかけて repair の回数が減っていることがわかる。

全体の流暢性(speed fluency)の値が向上した要因の一つとして、repairの回数と長さが少なく、短くなったことが考えられる。Repairは学習者が自身のスピーチをモニターし、発話の最中に自身の発話に誤りがあると感じた時に起こるものであり、モデルスピーチとの比較+筆記ランゲージングを通じて誤りと思う発話をする頻度が多少減った可能性がある。ある学習者は事前テストで、“(A boy painting) A boy is (painting) cutting paper.”と発話した(カッコ内の語句は repairされる前の誤りと学習者が気づいた表現)。この学習者は進行形に必要なbe動詞が抜けていたことと、色を塗っているわけではなく切っていることに気づき、動詞を paint から cut に変えた。この学習者は最終的には正しい発話を産出することができたが、それに到達するまでに repair を多く繰り返し、流暢性が低くなる発話をしてしまった。その後にモデルスピーチを聞き、筆記ランゲージングで気づいたことへの理解を深めたことで、直後テストでは “A boy is cutting some

paper.”と repair もなく発話することができた。遅延テストでは、“A boy (cutting) is cutting some paper.”と発話し、進行形の部分で repair したものの、若干の repair で済ませることができた。個別の事例ではあるが、他の学習者も似たような変化を示しており、モデルスピーチと筆記ランゲージングを用いた学習方法で流暢性の評価を下げる要因の一つである repair の頻度を減らすことができ、スピーチの質を高められることができると考えられる。

Q3についてもおおむね、Q2と同様の結果となった。Q3では、RQ1でも述べたように、学習者の多くが、so や because について気づき、それを自身のスピーチに取り入れることができたことから、一文当たりに発話するシラバブルの量が so や because を使用することで伸びたことも要因の一つとして考えらえる。接続詞を使用することで、節が増え、文全体が長くなつたということである。

モデルスピーチとの比較+筆記ランゲージングにより、学習者のスピーチはよりスムーズになり、スピーチの質を流暢性の面で高めることに貢献する可能性があることがわかった。

5.3 RQ3(スピーチの正確性の変化)について

Q2の正確性は、事前テストと直後テスト・遅延テスト間で有意に向上したという結果を得られた。全体的な結果としてはモデルスピーチとの比較+筆記ランゲージングにより正確性は向上する可能性があることが結果としてわかったが、以下では個別の事例を見ていき、留意すべき点を述べていく。

まず、語彙的正確性の側面をさらに詳細に見ていく。Q2の事前テストにおいて多くの学習者は適切な動詞を使うことができなかった。10名の学習者のうち、7名は “The man is pouring coffee” の「注ぐ」の意味である pour が言えなかった。この動詞は筆記ランゲージングの記述を見ると多くの学習者が知らなかった語であったことが考えられる。7名の学習者が pouring の存在には気づいたが、適切に表記できなかった。例えば「ポーリング」や「プアリング」、「ポール」のようにカタカナで表現するか、1人の学習者は “poling” のように書いた。受信語彙的な知識にも

入っていなかった可能性が高い。結果として、1回目に“A boy is take a coffee.”と発話した学習者は直後テストでもそれを繰り返し、また別の学習者は“A man spok?”(スボックのように発音した)coffee”と正確にpourを発話することができなかった。また、直後テストではpouringと発話できた学習者の中でも遅延テストでは3名が発話できなくなっていた。この結果の理由として、前述したが、pourという動詞を学習者は知らなかっただけの可能性がある。また、語彙知識にない語彙を使うことは負荷が大きかったことが考えられる。このことには実験デザインが影響していることも十分考えらえる。本研究では、モデルスピーチはリスニング媒体でしか学習者は接することなく、文字情報は与えられていない。音の情報だけでは語彙を使うまでには至らなかった学習者もいたということが考えられる。この結果からリスニングによるモデルスピーチだけでは未知の語彙を学ぶ上では不十分であることが考えられる。明示的に文字を示してあげるなど、文字情報も与えるような指導の工夫が必要かもしれない。もちろん、本研究では、文字によるインプットと音声によるインプットで気づきや習得が変わるとかを直接的に比較した研究ではないので、これら考察は予想の範疇を超えないという点は注意が必要である。

現在進行形に対しての気づきについて注目すべき事例がいくつかあった。ある学習者は事前テストで一度も正確に現在進行形を使用することができなかった。例えば“He is do my homework.”のようにV-ingができていなかったり、“The man wearing suits.”のようにbe動詞が抜けていたりした。この学習者はモデルスピーチとの比較では筆記ランゲージングに現在進行形に関しての記述が1つもなかった。その結果、直後テストでも同様にbe動詞かV-ingのどちらかが必ず抜けてしまっていた。この学習者はモデルスピーチを聞いても自身の使用した現在進行形が誤りであったことに気づけず、現在進行形に注意が向かなかったことが考えらえる。これもあくまで個別の事例ではあるが、モデルスピーチとの比較だけでは言語形式(現在進行形のような文法事項)に注意が向かず、スピーチの質向上させることができない場合も可能性としてあ

ることが分かる。こういった気づきに関する個人差についても、今後の研究でもっと注意深く見ていく必要があるだろう。

さらに、語彙に対して意識が集中し、できていたはずの現在進行形の形式上のエラーが生じた個別の事例もあった。事前テストにおいては5文すべてで現在進行形が適切に使用できていたものが、直後テストにおいては“The boy cutting some paper.”や“The girl is feed a cat.”のようにbe動詞とV-ingのどちらかが抜けている文が4文で、適切に表現できたものは1文のみであった。この学習者の筆記ランゲージングでの記述を見ると言語面については4件述べていた。具体的には「餌をあげる feed」、「下の男性 ポーリング some」、「女性 talking on the phone」、「男性 putting on the jacket」となっており、feedに関しては語彙情報にしか注意が向いておらず、be動詞にもV-ingにも注意が向いていない可能性がある。さらにそれ以外の3つについてはbe動詞を書いておらず、この側面に対して注意が向いていない可能性が考えらえる。

これらの現象はHanaoka(2007)で論じられているScope of Noticingと関連している。Hanaokaの研究では、「渋滞に合う (be caught in a traffic jam)」を表現する際にtraffic jamという表現がはじめのライティングでできなかった学習者はモデル文との比較でtraffic jamは使用できたが、be caught in a traffic jamというような、traffic jamの前にくるコロケーションまでは正確に使用することができず、“Because it is traffic jam in the morning”というように文法的に誤った表現をした。しかし、traffic jamという表現ははじめからできていた(知っていた)別の学習者はbe in the traffic jamを自身のライティングに取り入れ、“He was caught in a traffic jam.”と正確に表現できていた。上記であげた学習者もHanaokaの研究であげられていた学習者と同様に、はじめのスピーチでfeedもpourも表現できていおらず、モデルスピーチとの比較でそれらの語彙に対しては気づき、自身のスピーチに取り入れることができたが、be feeding V-ingというコロケーション情報にまでは気づきが及ばず、もともと正確に使用していた現在進行形に形式上のエラーが生じた可能性がある。また、

この学習者はモデルスピーチとの比較の中で、語彙情報に認知資源を向け過ぎて、文法的側面に注意が行かなかった可能性もある。直後テストでは、語彙は結果として、適切に使用できたが、現在進行形の形式上のエラーが生じたことも考えられる。また、遅延テストでは1文を除き、現在進行形ができていた。時間が経つにつれて、語彙知識を自身の知識として取り入れることができ、その分の認知資源を文法的側面に充てることができたことも可能性として考えられる。

Q3に対しても事前テストと直後テストの間で行われた介入で正確性の向上が見られ、ある程度その向上を維持できたことが考えられる。特にsoやbecauseを使用できるようになった学習者が多く、モデルスピーチとの比較+筆記ランゲージングを通じて、soやbecauseの接続詞に気づくことができ、それを自身のスピーチに取り入れることができるようになったことが考えられる。RQ1に対しての考察でも述べたが、多くの学習者の筆記ランゲージングの記述の多くにsoやbecauseの接続詞に対しての気づきが表出され、学習者は実際に自身のスピーチに取り入れることができた可能性が考えられる。しかし、遅延テストで、不正確に接続詞を使用した学習者のうち、1名の発話は“A dog is dirty because a girl want to wash dog.”のように、becauseの前後の節が逆になるという典型的な誤りをしている。接続詞についてモデルスピーチを聞いただけでは正確に理解できなかった可能性が考えられる。学習者の中には、接続詞の意味を明示的に教えることで理解が深まる学習者がいることも考えられる。

6

結論と実践方法の提案

まず、本研究の結果をまとめると、RQ1については、気づきの多くは語彙情報についてであり、多くの学習者は自分のスピーチとモデルスピーチの比較を通じてモデルスピーチの語彙的特徴に注意が向いていたことがわかった。また、Q2において、学習者はどの程度の気づきかはわからないが、現在進行形に注意を向けていた可能性も

高かった。さらに、Q3のタスクにおいてはsoやbecauseと言った接続詞の文法事項にも注意が向き、意味機能の高い接続詞の卓立性が高いことが理由であると考察した。

RQ2については、モデルスピーチとの比較+筆記ランゲージングによって学習者のスピーチの流暢性が向上し、そして時間が経っても流暢性は維持される可能性があることがわかった。Repair(言い直し)の回数と長さが少なく短くなつたことや、soやbecauseなどの接続詞を使用するようになって文全体が長くなったことが流暢性の向上に寄与したと考察した。この学習方法で流暢性の観点でスピーチの質は向上することが考えられる。

RQ3についても、モデルスピーチとの比較+筆記ランゲージングで学習者のスピーチの正確性が向上し、時間が経っても正確性は維持される可能性があることがわかった。全体的に語彙的正確性は向上し、維持されたが、注意すべき事象があった。学習者がスピーチ時には知らなかった語については、モデルスピーチとの比較だけでは気づくことができず、インテイクまで至らなかつたことが考えられるケースがあった。さらに、現在進行形が事前テストで全くできなかった学習者がモデルスピーチとの比較から現在進行形の形式に気づけなかったケースがあった。最後に、事前テストでは現在進行形を正確に発話できていた学習者でも、語彙情報に焦点が向き、これまで見られなかった現在進行形の形式上のエラーが表れた学習者もあり、その結果からHanaoka (2007) のいう Scope of Noticing が関与している可能性や、学習者の意識が語彙に向いた結果、文法に認知資源が向かなかった可能性があると考察した。

これらの結果と考察を踏まえると、スピーキングの指導に対していくつかの提案ができる。1つ目に、学習者が受信語彙として語彙知識にある語で1回目のスピーチで発信できなかった語については、モデルスピーチを聞かせることで学習者自身に気づく機会を与え、筆記ランゲージングなどをさせた上で再度スピーチさせる。その際に語彙の誤りが含まれていた場合には、直接的に明示的指導として「〇〇という日本語は〇〇という英語で表現できる」というようにフィードバック

を与えるようにするという指導手順である。2つ目に、現在進行形や接続詞などといった文法事項にも語彙的特徴ほどではないが注意が向く可能性があることを踏まえると、現在進行形や接続詞などの意味的・視覚的に卓立性の高い文法事項も同様の方法で指導することもできるかもしれないという点である。最後に、モデルスピーチとの比較+筆記ランゲージングによって、いくつかの制約があるが、スピーチの質(流暢性と正確性)は向上することが考えられるので、学習者が自主的にスピーキングの練習をする際には、本研究で用いた【ステップ1】学習者自身によるスピーチ、【ステップ2】モデルスピーチとの比較+筆記ランゲージング、【ステップ3】再度スピーチという指導・学習方法を学習者に提示することを提案したい。また、スピーキング指導をクラス全体に対して一斉に行う際にもこの方法を応用させることができることも併せて提案したい。

7

今後の展望

本研究はモデルスピーチとの比較+筆記ランゲージングの効果を少人数ではあるが、参加した学習者にとっては有効な手立てであることがわかった。ただ、本研究はより大きな研究のパイロットスタディとして位置づけて行ったものであるので、結果の妥当性や考察したものは個別の事例の域を超えないことは留意点として述べておきたい。今後はより規模の大きく妥当性のある研究にするためにいくつか注意しなくてはならない点がある。1点目に、本研究では筆記ランゲージングで学習者の気づきを見たが、他の気づきや注意を扱った研究で使用されている Stimulated Recall(刺激思考発話) や Eye-tracking(視線計測)などの手法も行い、タスクを行っている最中の学習者の気づきや意識も見ていく必要がある。2点目としては、より長いスピーチにおいても同じ指導方法が効果的か見ていく必要がある。本研究のタスクは短いものであり、学習者がモデルスピーチを聞いた後に行った発話はモデルスピーチをただ暗記しただけであった可能性もある。さらに、より長いスピーチにすることで、発話の

複雑性も見ることができる。パフォーマンスの側面として、流暢性・正確性に加え、複雑性の観点もあり、学習者のスピーチパフォーマンスの複雑性の側面も考察することで、パフォーマンスの変容がより詳細に評価することができる。最後の3点目は実験デザインについてである。本研究には統制群を設定しておらず、モデルスピーチとの比較+筆記ランゲージングだけがスピーチの質の変容に影響を与えたのか、それともタスクを繰り返すだけ、つまりモデルスピーチとの比較や筆記ランゲージングなしで、同じスピーキングタスクを再度行うことでスピーチの変容が起きたのかが本研究のデザインではわからない。より綿密な実験デザインにすることでより正確なデータによる分析を行うことができる。今後は、本研究でわかったこと、以上で述べた点を踏まえて、実験を続けていく。

謝辞

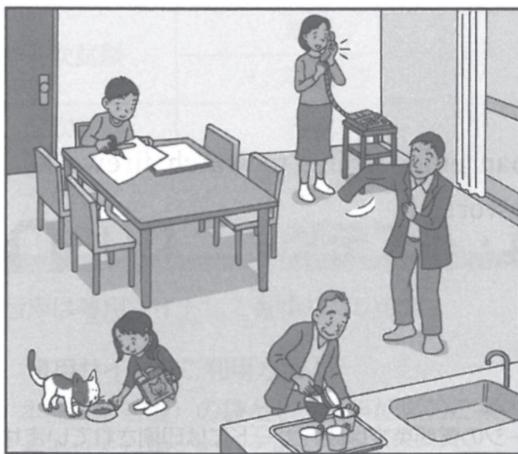
本研究の実施・発表にあたり、貴重な機会を与えてくださった公益財団法人 日本英語検定協会の皆様、ならびに選考委員の先生方に心から御礼申し上げます。特に助言者として、論文を丁寧に見ていただき、大変有益なご助言をくださった和泉伸一先生には深く感謝いたします。最後に、本研究に協力して下さった申請時の勤務校であった小倉南高等学校の生徒の皆さんにはこの場をお借りして感謝申し上げます。

参考文献(*は引用文献)

- * 和泉伸一(2016a)「フォーカス・オン・フォームとCLILの英語授業」東京:アルク
- * 和泉伸一(2016b)「第2言語習得と母語習得から「言葉の学び」を考える」東京:アルク
- * 文部科学省.(2018). 平成29年度英語教育実施状況調査について 結果概要 高等学校
(https://www.mext.go.jp/component/a_menu/education/detail/_icsFiles/afieldfile/2018/04/06/1403469_02.pdf)
- * 文部科学省.(2019). 平成31年度(令和元年度)全国学力・学習状況調査 中学校英語「話すこと」調査 検証報告書
(https://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/afieldfile/2019/09/20/1421443_02.pdf)
- * Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12(3), 267-296.
- * ELAN (Version 5.2) [Computer software]. (2018, April 04). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from <https://tla.mpi.nl/tools/tla-tools/elan/>
- * Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305-352.
- * Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, 35(3), 407-432.
- * Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, 20(3), 321-340.
- * Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 459-479.
- * Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, 21(4), 332-347.
- * Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461-473.
- * Ishikawa, M. (2018). Written languaging, learners' proficiency levels and L2 grammar learning. *System*, 74, 50-61.
- * Ishikawa, M., & Suzuki, W. (2016). The effect of written languaging on learning the hypothetical conditional in English. *System*, 58, 97-111.
- * Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24(4), 541-577.
- * Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24(2), 168-196.
- * Izumi, S. (2013). Noticing and L2 Development: Theoretical, empirical, and pedagogical issues. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka, (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp.25-38). Honolulu: University of Hawaii, National Foreign Language Resource Center.
- * Kahng, J. M. (2014). *Exploring the production and perception of second language fluency: Utterance, cognitive, and perceived fluency*. Michigan State University. Second Language Studies-Doctor of Philosophy.
- * Kang, E. (2020). Using model text as a form of feedback in L2 writing. *System*, 89(1), 103-196.
- * Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge.
- * Schmidt, R. W. (1990). The role of consciousness in second language learning1. *Applied Linguistics*, 11(2), 129-158.
- * Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. *Implicit and Explicit Learning of Languages*, 22, 165-209.
- * Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- * Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, 64(1), 103-131.
- * Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(1), 1-14.
- * Storch, N. (2008). Metatalk in a pair work activity: Level of engagement and implications for language development. *Language Awareness*, 17(2), 95-114.
- * Swain, M. (1995). Three functions of output in second language learning. In G. Cook and B. Seidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honor of HG Widdowson*. (pp. 125-144). Oxford: Oxford University Press.
- * Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. *Sociocultural Theory and Second Language Learning*, 97, 114
- * Swain, M. (2006). Languaging, agency and collaboration in second language learning. *Advanced Language Learning: The Contribution of Halliday and Vygotsky*, 95-108.
- * Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82(3), 320-337.
- * Tavakoli, P., & Skehan, P. (2005). 9. Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-273). Amsterdam: John Benjamins.
- * Tavakoli, P., Nakatsuhara, F. and Hunter, A. (2020), Aspects of Fluency Across Assessed Levels of Speaking Proficiency. *The Modern Language Journal*, 104(1), 169-191.
- * VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten and J. Williams (Eds.). *Theories in Second Language Acquisition* (pp. 115-135). Mahwah: Laurence Erlbaum.

資料1:

Appendix A: Q2 マテリアル



Appendix B: Q3 マテリアル



Instructions:

Q2: Now, please look at the people in Picture A. They are doing different things.

Tell me as much as you can about what they are doing.

Q3: Now, look at the girl in Picture B. Please describe the situation.

Model answers:

Q2: A boy is cutting some paper. / A girl is feeding a cat. / A man is pouring some coffee. /

A woman is talking on the phone. / A man is putting on [taking off] his jacket.

Q3: Her dog is dirty, so she is going to wash it. / She is going to wash her dog,

because it is dirty.

第32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

AIを活用したライティング能力の育成

— CEFR B1 から B2 へのレベルアップを狙って —

研究者:新潟県／新潟県立 津南中等教育学校 教諭 松井 市子

《研究助言者:小泉 利恵》

概要

本研究ではライティング力をCEFR AからBレベルへ上げることを狙った指導の転換を目指し、人工知能(AI)によるフィードバック機能を持つWrite & Improveを活用した指導の効果を検討した。その結果、A2レベルの生徒において、AI feedbackが有効であることが分かった。B1からB2へのレベルアップは外部指標では見られなかったが、定期考査に位置づけた評価タスクでは見られ、活動を継続することで次回の外部指標で示される可能性が高いことが分かった。AIを活用すると、スペルミスなどは生徒自ら修正でき、読みやすい出来栄えになった。また、AIが出すCEFR判定の結果が思ったより良いことに励まされた生徒も多かった。日本の高校生は「話す」「書く」のアウトプット技能の育成に課題があることが指摘されているが、ICT環境や教材を活用すれば自立学習を促進することができる。多くの時間を割いてきた文法学習や「読む」「聞く」のインプット活動を見直すことになり、結果的にアウトプット活動の時間を確保できることにつながる。今後は、浮いた時間でより生徒同士や教師との「やりとり」を楽しみ、「話す」「書く」のアウトプット技能を育成しながら、読み応え、聞き応えのある教材を生徒と共に味わいたい。

1

はじめに

新学習指導要領(文部科学省、2017)では、「社会に開かれた教育課程(カリキュラム・マネジメント)」の実現を中核に、「何が出来るようになるか」「何を学ぶか」「どのように学ぶか」の3つの視点で改訂の方向性が示された。特に、①知識及び技能、②思考力、判断力、表現力等、③学びに向かう力、④人間性等の育成を、「主体的・対話的で深い学び(アクティブ・ラーニング)」で、質的改善に留意しながら進めていくことが強調された。また、経済協力開発機構(OECD)でも、「教育とスキルの未来 2030事業(Education 2030)」が示され、複雑で予測の難しい2030年の世界を生きる子供たちに育成すべき資質・能力(コンピテンシー)が何で、どのように各国で協力すべきかが示された。特に、第1フェーズ(～2018年)では21世紀向けのコンピテンシーが提案され、第2フェーズ(2019年～)ではコンピテンシー育成のための学習環境や教育システムの提案が予定されている。

このような流れを受けて、これから時代を生き抜く子供たちに育成すべき英語力、およびその育成方法や教育環境を念頭に、本研究では人工知能(AI)を活用したライティング指導を取り上げる。大学入試改革の一環として取り入れられた民間英語検定試験の活用も本研究に関連している。CEFR(外国語の学習・教授・評価のためのヨー

ロッパ言語共通参照枠)は外国語運用力をA1～C2の6段階で示しているが、この指標が大学入試の選抜基準や民間英語検定試験でも使用されるようになった。2018年度英語力調査結果(高校3年生対象)によると、A2以上の割合が「聞くこと33.6%」「話すこと12.9%」「読むこと33.5%」「書くこと19.7%」で、4技能のバランスに課題がある上、目標50%にはどの技能も達していない。特に、「話す」「書く」のアウトプット技能の育成に課題があり、Aレベル「基礎段階の言語使用者」からBレベル「自立した言語使用者」の育成を見据えた指導法の確立が求められている。しかし、アウトプット技能の育成はその評価の信頼性や妥当性、クラスサイズが原因の実行可能性が問題視され、なかなか進まないのが現状である。本研究では、コンピューターによる自動採点評価システム(AI)を用いたライティング指導を扱う。AIの活用でライティング技能の育成に効果が見られれば、アウトプット技能の育成をねらった評価や活動の実行可能性が高まる期待する。

ライティング研究におけるAIや自動採点フィードバックについて、学習者の自立を促したり学習意欲を向上させたりするツールとしての効果は指摘されている(齋藤, 2017; Brown, 2018)。また、流ちょうさの伸長も見られる(Ohta, 2008)。一方、どのレベルの生徒にもAIフィードバックが効果的かどうかの検証はあまり行われていない。本研究では、CEFR AからBレベルへの指導の転換、およびB1からB2へのレベルアップを狙ったAIを用いたライティング指導の効果を報告する。

2 実践の背景

2.1 CEFR-J

本校は県立中高一貫校で、英語力に関してはCEFR-J(cefr-j.org)に基づくCan-doリストで到達目標を1年A1.1～6年B2.1に設定している(資料1)。4年生(高校1年生)の目標はB1.1で、ライティングはB1.1～1.2のタスクを教科書や学校行事に関連させて作成する。

2.2 ICT環境とその活用

本校には、県から供給されるネット回線とパソコン約40台がパソコン教室に完備されている。これらを使用して本実践を遂行することも可能だが、Write & Improveを使用するにあたって県への毎月の申請が必要だったり、Internet Explorerではうまく起動しなかったりで、不便に感じることがあった。そんな中、新入試対応の一環で、2018年度半ばに学校の後援会予算(4年契約200万円)で、学校独自回線のWi-Fiを完備させた部屋にタブレットパソコンを40台設置した。タブレットパソコンはキャビネット収納のためセッティングに手間がかかるが、Write & Improveの使用やその成果物の提出先であるedmodoの使用においても便利だと感じ、ライティング活動はWi-Fi教室で行った。2019年度半ばにはPC教室のパソコンの入れ替えで旧型が処分対象になったため、これも後援会予算(約100万円)でバージョンアップして再利用し、第2PC教室を設けて学校独自回線環境も拡張した。本研究はこの第2PC教室での実践がメインとなる。

生徒の家庭のICT環境調査を行った結果、2018年度後半は54名中3名がインターネットを使用できない、もしくは使用できるデバイスを所持していない環境であった。前期生(中学1～3年生)は携帯端末所持禁止という校則も関係していると考えられる。しかし、本実践を行った2019年度後半では全員使用できる環境であった。データ容量や使用制限で活動や課題に家庭で取り組めない生徒があり、そのような生徒を対象に、週1回第2PC教室を放課後開放した。

ライティング力の育成には大量のインプットはもとより、豊富な語彙と基本的な文法知識が欠かせない。授業ではアウトプット活動に多くの時間を割けるよう、インプットは家庭で取り組むよう生徒を促し、反転学習を目指した。語彙力育成のため『DataBase 第5版4500』(桐原書店)を3年2月から使用し、小テストを毎木曜日に実施し、定期考査にも出題した。生徒は付属のアプリケーションをダウンロードしてインターネットで演習問題や音読に取り組むことができる。考査前1ヶ月は5分帯学習としてペア活動で語彙の定着を図った。文法力育成のため『5 STAGE 英文法完

成 Book1～3』(数研出版)を3年12月まで、以降は『ジーニアス総合英語』(大修館書店)を使用した。また、基礎力を診断する「到達度テスト」に付属の学習アプリ「スタディサプリ」(リクルート)を3年6月から使用した。『ジーニアス総合英語』と「スタディサプリ」は文法説明をインターネット動画で視聴できる。「スタディサプリ」は読解・資格対策・語彙などの演習問題や到達度テスト関連問題も教師がインターネットで課題配信できる。4年6月までは文法の家庭学習用週課題を紙媒体で課し、提出チェックをして、未提出者への声掛けや補習指導を行ったが、6月以降は廃止した。朝テスト(10分間)は全校での取り組みで、8割未満の理解度の生徒には放課後補習を実施する。英語は週1回英文法について出題した。参考書やアプリを使用して家庭で予習し、朝テストで定着度を確認する、という流れで、授業では文法学習のためだけの時間をほとんど設けなかった。

リスニング力育成のため学習アプリ「スタディサプリ English」(リクルート)の日常会話コースを3年12月から使用している。導入当初は週課題として取り入れ、授業で週1回、ALT授業時のパフォーマンスタスク待機時に使用したが、4年6月からは家庭学習として使用することを促している。さらに教科書に付属の学習アプリ「English Central」のサービスが始まったため、家庭学習用に導入した。教科書の内容が動画で視聴でき、単語演習や音読とその録音、AIによる採点機能も付いている。この導入がきっかけで授業の指導方法が大きく変わり、インプットとアウトプット活動の割合がほぼ1対1となった。

2.3 インターネット上の無料システムの利用

2.3.1 Write & Improve

(AIによるライティングのフィードバック)

ライティングの校正には、主に Write & Improve というインターネット上のサービスを使用した。2016年度から始まったサービスで、ケンブリッジ大学出版とケンブリッジ大学英語検定機構がケンブリッジ大学と共同開発した AI による自動フィードバックシステムである。ケンブリッジ英検や IELTS などの民間英語検定試験対策になるライティングタスクの他、レベル別

に様々なタスクが用意されている。書いたものを AI が瞬時に CEFR 判定や(正しい解答は提示されないが)エラーチェックをしてくれる。本研究では無料版を使用した。まず、教師としてアカウントを作成し、生徒用の “workbook(タスク)” を作る。次に、workbook の専用 code を生徒に知らせ、アカウントを作らせる。アカウントがなくてもサービスを使用できるが、アカウントがあると生徒は教師作成の workbook に取り組めたり、過去のタスク履歴を参照したりできるためアカウントは作らせた。

2.3.2 Text Analyzer

モデル文の CEFR レベル判定には CEFR-J サイト内の Text Analyzer (cefr-j.org) や CVLA (CEFR-based Vocabulary Level Analyzer) Ver1.1 (Uchida & Negishi, 2018) を使用した。

2.3.3 edmodo

edmodo は教育目的の非同期型プラットフォームである。教師と生徒、保護者がそのプラットフォームを活用して、メッセージのやり取りやテキスト・音声・動画の投稿をしたり、生徒に課題や小テストを提出させたりすることが出来る。

ライティングタスクは ICT(情報通信技術)スキル向上を狙って、Word ドキュメントで edmodo 上に課題として提出するよう促した。Write & Improve でタスクに取り組み、その内容を Word に貼り付け、edmodo に提出するという流れである。

3 実践方法

3.1 目的

本研究の目的は AI (Write & Improve) によるフィードバックが生徒のライティング力育成に効果をもたらすかどうかについて、以下の仮説を設定し検証することである。

仮説1 B1の生徒は AI によるフィードバックで B2 にレベルアップできる

仮説2

A2の生徒はAIによるフィードバックでB1にレベルアップできる

仮説3

A1の生徒はAIによるフィードバックでA2にレベルアップできる

3.2 対象

本校4年生(高校1年生)2クラス54名を対象に、Write & Improve(以後、W&I)を活用したライティング指導の効果を、コミュニケーション英語I(4単位)の授業で、『Pro-Vision English Communication II』(桐原書店)のLesson 1~5を使って7~12月の夏休みと海外研修準備期間を除く実質3か月間調査した。生徒の希望に基づき各クラスで発展クラスと基礎クラスを編成した。発展クラスは、教科書英文をモデル文としてB2レベルの英文が書けることを目標とし、教科書を概ね自力で理解しようとする意欲があり、家庭学習が習慣化している生徒が選んだ。基礎クラスは、教科書付属のサマリー英文をモデル文としてA2レベルの英文が書けることを目標とし、教師の助けありの教科書理解を求め、家庭学習の習慣化が困難な生徒が選んだ。学年2クラス(各クラス27名)のうち、希望に基づくクラス分けの割合は発展3:基礎1であった。本研究実践者A(教職歴20年)が1組の基

礎クラスと2組の発展クラスを、教師B(教職歴7年)が1組の発展クラスと2組の基礎クラスを担当した。

3.3 手順

指導は学習タスク、練習タスク、評価タスクの3段階で行った。まず、CEFRの各レベルのライティングがどう違うのかを学ぶために学習タスクを4つ示した。1つは英検準2級、2級、準1級(2018年第2回)のライティングセクションの模範解答、もう1つはGTEC付属の「自由英作文の書き方演習冊子」のLevel 2~5の参考英作文例であった。タスクはすべて意見文で、トピックは英検準2級“Do you think it is better for students to study alone or in a group?”(50~60語)、2級“Today, some people buy products that are good for the environment. Do you think buying such products will become more common in the future?”(80~100語)、準1級“Should Japan do more to protect its historic sites?”(120~150語)、GTEC「音楽の良い点について述べよ」であった。W&IやText Analyzer(TA)を使用してCEFRを判定すると結果は表1の通りであった。学習タスクはモデル英文を比較させ、それぞれの特性を学ばせることが狙いであった。

■表1: 学習タスクのCEFRレベル

	英検準2	英検2	英検準1	GTEC2	GTEC3	GTEC4	GTEC5
Write & Improve	A2	B1	B1	A1	A1	A2	A2
Text Analyzer	A2	B1	C1	A2	B1	B1	B1

練習タスクは教科書を使用した。学習タスクで学んだことを活用して英語運用力を高めることが練習タスクの狙いである。手順は表2の通りで、活動①~④は第2PC教室で合同実施し、活動⑤は数単元分まとめて発展・基礎クラス毎に分かれて実施した。

教科書は各単元4パート構成だが、長い英文に慣れさせるため4パート分まとめて活動した。1コマ目の読み取り活動①は700~900語程度の英文を一気に読み、読むのにかかった時間をパソコンのアラーム機能を使って計測した。その後、

英文中の教科書のQuestionの答えの箇所に下線を引き、かかった時間と正答率でwpm(語数÷読むのにかかった秒数×60秒)を計算した。また、活動①と③の間を1週間空け、生徒にはサマリー作成に向けて教科書付属の学習アプリ「English Central」を使用して自宅学習を進めるよう指示した。2コマ目の聞き取り活動②では、教科書の英語メモ“Notes”(Q&A形式)を埋められるまでEnglishCentralを繰り返し聞いた。聞き取りが得意な生徒はアプリの語彙問題等に進んだ。活動②と③の間に授業がある場合は、発展・基礎クラ

ス毎に『PRO-VISION English Communication I』に戻って速読やポイント文の翻訳をした。

■表2: 練習タスクの活動手順

1コマ48分	発展クラス	基礎クラス
① 1コマ	単元の読み取り(Q&A)	
② 1コマ	単元の聞き取り(英語メモ作成)	
③ 1~2コマ	英語メモを元に単元サマリー作成 → Write & Improve で校正 → 提出	単元のモデルサマリーのタイピング → 消して自力で復元 → Write & Improve で校正 → 提出
④ 1~2コマ	単元 Speaking パフォーマンステストとフィードバック	
⑤ 1コマ	単元のまとめ(リスニングと文法表現の確認)	

活動③は発展・基礎クラス毎に異なるタスクを指示した(図1)。発展クラスは活動②で作成した英語メモを元にライティングサマリーを目標15分で作成し、残り時間はW&Iで校正し、最後にWordドキュメントに校正前後の英文・語数・CEFR・作成時間を記入してedmodoに提出した。基礎クラスは1回目にモデルサマリーをW&Iに

タイピングし、2回目にそれを消して活動②で作成した英語メモを元に自力で復元し、モデルと自作の英文・語数・CEFR・作成時間をWordドキュメントに記入してedmodoに提出した。基礎クラスのタイピングタスクは、12月に受験するリンクガスキル(CBT:コンピューターベースのテスト)に対応させるためでもあった。

■図1: 練習タスクの完成イメージ(発展・基礎クラス共通)

■表3: 練習タスクの語数とCEFRレベル

	Lesson 1	Lesson 2	Lesson 3	Lesson 4	Lesson 5
本文語数	715	708	811	831	905
Text Analyzer	B1	B2	C1	B2	B2
モデルサマリー語数	172	178	163	122	186
Write & Improve	B2	B2	B1	B2	B2
Text Analyzer	B2	C1	C1	B1	C1

表3は練習タスクに使用した教科書本文と付属のモデルサマリーの語数とCEFRレベルである。W&Iは600語以上の英文のCEFR判定に対応しないため、本文のCEFRはText Analyzerのみ示した。

活動④は、単元についてのペアでの5分間フリートークで、ALT(外国人指導助手)と2名のJTE(日本人教師)で分担し、5単元分同じペアを評価した。

5分間ほぼよどみなく単元に関するやり取りができるか、3つ以上の質問ができるか、の2点について、両方できてA、いずれかでB、いずれもできなければCの評価を与えた。生徒がやり取りにおいて行き詰まつたり語法や文法に誤りが見られた場合、教師は介入したりrecast(誤り訂正)したりした。生徒は待機中、『スタディサプリ English』などの

Can-do Description [Writing B1.1]					
I can write a description of substantial length about events taking place in my immediate environment (e.g. school, workplace, local area), using familiar vocabulary and grammar.					
Rubric [Writing B1.1]					
Criteria 規準		A	B	C	Total score
Content 内容	Summary 聞いたり読んだりした内容の要点について述べられている	詳しく述べられている： 3点	最低限述べられている： 2点	無関連に述べられている： 1点	/3
	Impression / Opinion 感想や意見等が述べられている	詳しく述べられている： 3点	最低限述べられている： 2点	無関連に述べられている： 1点	/3
Structure 構成	Cohesion & Coherence ①代名詞や接続詞などを効果的に使用し②分かりやすい流れと構成である	①②とも効果的にある： 2点	いずれか意識している： 1点	いずれも無し：0点	/2
	Vocabulary & Grammar 読み手や状況を考えた日常生活語彙や文法を適切に使用している	ほど適切に使用している： 2点	いずれか使用している： 1点	いずれも無し：0点	/2
				/10	

■図2: 評価タスクB1.1のRubric

Can-do Description [Writing B1.2]					
I can write narratives (e.g. travel diaries, personal histories, personal anecdotes) in several paragraphs, following the order of events. I can write personal letters which report recent events in some detail.					
Rubric [Writing B1.2]					
Criteria 規準		A	B	C	Total score
Content	Time & Place & Occasion 場面や状況について記述している	詳しく述べられている： 3点	最低限述べられている： 2点	述べられない：0点	/3
	Development 出来事の順序と展開についての複数段落による記述をしている	詳しく述べられている： 3点	最低限述べられている： 2点	展開されない：0点	/3
Language 言語	Vocabulary 読み手や状況を考えた語彙を適切に使用している	正確：2点	だいたい正確：1点	誤りが多い：0点	/2
	Grammar 読み手や状況を考えた文法を適切に使用している	正確：2点	だいたい正確：1点	誤りが多い：0点	/2
				/10	

■図3: 評価タスクB1.2のRubric

学習支援アプリを使用して自習した。

活動⑤は、定期考査前に数単元分まとめて、発展・基礎クラス毎に単元末のリスニングと文法演習、ポイント文の翻訳活動を行った。

評価タスクは定期考査に位置づけ、練習タスクの定着度を測ったり、外部指標との比較に利用したりした。定期考査はコミュニケーション英語Ⅰ・英語表現Ⅰ合同で110分200点満点。ライティングセクション60点のうち、20点分はB1.1(説明文や意見文)とB1.2(物語文)の自由記述問題を課した。図2・3のRubricを学習タスク導入時に提示し、評価タスク時まで使用した。B1.1とB1.2のモデル文は『Pro-Vision English Communication II』に掲載されている説明文、意見文、物語文を利用し、それぞれの構成や流れの違いを学習した。説明文形式には本研究の練習タスクに取り組むことで慣れていくが、意見文や物語文は教科書にあまり掲載されていない。そのため、発展クラスでは意見文や物語文をディベートやディスカッション、朗読活動後に練習タスクとして授業で1回扱い、添削して返却した。評価タスクでは基礎クラスはB1.1(説明文)を、発展クラスは外部指標CEFR Aレベルの生徒はB1.1(意見文)を、Bレベルの生徒はB1.2(物語文)を選択するよう指示した。トピックは「留学生との交流授業」や「海外研修」など、異文化体験が絡むイベントを扱った。採点は英語表現Ⅰの授業担当者3名(発展クラス担当の教師A=本研究実践者と教師C=教職歴6年;基礎クラス担当の非常勤講師D)が行った。100語未満の記述に関しては、その量に応じて内容・構成、もしくは語彙・文法語法のいずれかのみ採点した。Rubricに基づくスコアの他、発展クラスには間接的フィードバック(間違えに下線を引く)を、基礎クラスには直接的フィードバックを与えた。また、基礎クラスには定期考査前の授業で事前添削も行った。解答用紙返却時、生徒はW&Iで校正し、CEFRレベルを判定した。

各科目評点100点中、定期考査50点、パフォーマンス(コミュニケーション英語はスピーキング、英語表現はライティングが中心のタスク)20点、課題10~20点、小テスト等10~20点のうち、練習タスクはコミュニケーション英語Ⅰの課題5点分(1単元1点)、評価タスクは英語表現Ⅰの定

期考査10点分(各考査のライティング自由記述問題平均点相当)に充てた。

4 結果と考察

4.1 分析対象データ

使用する分析対象データは以下の通りである。外部指標の民間英語検定試験に関しては、同じ生徒の3年と4年時のデータを使用した。データの分析にはJASP Version 0.11.1.0とjs-STAR Version 9.8.3jを使用した。

【外部指標】

進研(B)模試7月・11月、全統(K)模試5月・8月
GTEC33B(3年①6月), GTEC36A(3年②12月),
GTEC37A(4年①7月), リンガスキル(4年②12月)

【内部指標】

練習タスク(Lesson 1~Lesson 5)のCEFR
判定および校正前後の伸び
評価タスク(1学期中間6月, 1学期末9月, 2学期
中間12月, 2学期末2月)の点数

なお、模試や練習タスクと評価タスクの難易度の調整はされていないため、スコアの違いはタスクの違いにも影響される可能性がある。本研究では大まかなパターンを把握する目的で用いる。また、民間英語検定試験の信頼性が保証されているという前提で分析を行う。

【事後アンケート】

家庭のICT環境、家庭の英語学習状況、英語学習への関心意欲態度
顕著な伸びが見られた生徒・伸びが見られなかった生徒の自由記述回答

4.2 データ分析結果

4.2.1 外部指標(模試)が示す対象者の英語力の変化

模試は高校段階での基礎学力や大学入学志望者

の相対的学力を判定することが目的である。本研究では模試の得点アップをねらった指導をしたわけではないが、CEFRのレベルアップをねらった指導が模試の得点にどのように反映されたかを検討する。

表4は全統模試5月(K5)と8月(K8)、進研模試7月(B7)と11月(B11)の英語全体とライティングの記述統計、図4・5はそのヒストグラムである。模試は記述式で、ライティングは並べ替え、場面設定作文、日本文英訳の形式が入っていた。クラスごとの両模試の英語全体とライティングで1回目と2回目のスコアに差があったかどうかを調べるために、まず両模試の正規性についてJASPを使ってShapiro-Wilk検定を行ったところ、正規性が認められた。そこで、Levene検定を行ったところ、2回目の進研模試ライティングで等分散

性が認められなかった($F = 7.108, p = .010$)。当模試のスコアは Welch の t 検定を、他は Student の t 検定を行った結果、英語全体に関して両クラスとも全統模試において有意に2回目のスコアが下がった(発展クラス $t(39) = 5.139, p < .001, d = 0.813$ 、基礎クラス $t(13) = 5.387, p < .001, d = 1.440$)。また、ライティングに関して基礎クラスの進研模試以外において有意に2回目のスコアが下がった(発展クラス K 模試 $t(39) = 3.516, p = .001, d = 0.556$ 、B 模試 $t(39) = 3.158, p = .003, d = 0.499$ 、基礎クラス K 模試 $t(13) = 3.211, p = .007, d = 0.858$ 、B 模試 $t(13) = 1.060, p = .308, d = 0.283$)。以上の結果から、CEFRのレベルアップをねらった指導は必ずしも模試のスコアアップにつながるとは限らないことが分かった。

■表4: 進研(B)模試と全統(K)模試の英語全体とライティングの結果(N=54)

		英語全体				ライティング			
		K5	K8	B7	B11	K5W	K8W	B7W	B11W
満点		200	200	100	100	30	30	20	20
平均	発展	119.10	104.26	61.00	58.20	20.33	17.33	12.90	11.13
	基礎	72.79	53.86	27.50	28.36	11.14	6.64	5.64	4.64
標準偏差	発展	23.99	28.22	14.10	14.67	5.30	6.73	2.86	2.86
	基礎	24.54	17.16	14.88	14.95	7.83	5.24	4.47	4.11
標準誤差	発展	3.79	4.46	2.23	2.32	0.84	1.06	0.45	0.45
	基礎	6.56	4.59	3.98	4.00	2.09	1.40	1.19	1.10
変動係数	発展	0.20	0.27	0.23	0.25	0.26	0.39	0.22	0.26
	基礎	0.34	0.32	0.54	0.53	0.70	0.79	0.79	0.89
最大値	発展	163	166	89	91	30	30	18	17
	基礎	112	85	58	49	24	14	17	11
最小値	発展	62	66	31	26	8	0	6	5
	基礎	33	30	0	0	0	0	0	0

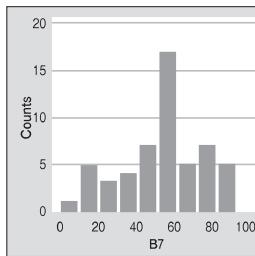
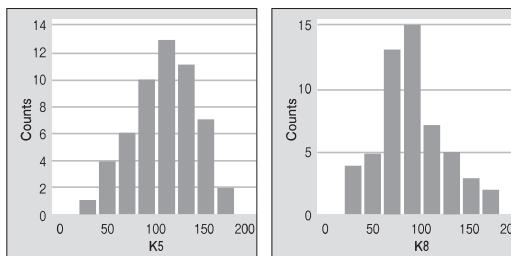
注:K5 = 全統模試 5月, B7 = 進研模試 7月, W = ライティング

4.2.2 外部指標(民間英語検定試験)が示す

対象者のライティング力の変化

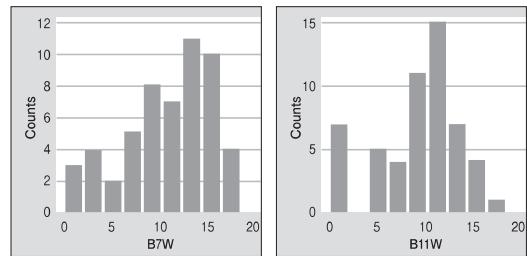
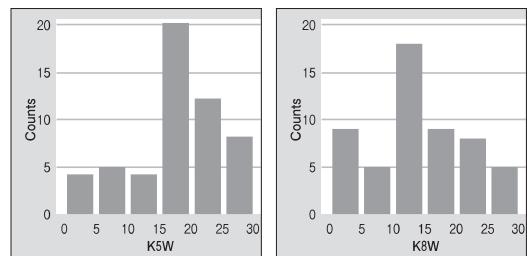
表5は民間英語検定試験(GTECとリンガススキル)を4回(3年①②、4年①②)とも受験した45名分のライティングのCEFR別人数を、図6はその割合の推移を示したものである。js-STARを使って χ^2 検定を行った結果、人数の偏りは有意であった($\chi^2(6) = 18.752, p < .01$)。残差分析によると

(表6)、3年①のA1と4年②のB1が有意に多く、4年②のA2が有意に少なかった。3年②と4年①では中学校の教科書から高校の教科書に変わり、指導方法や活動内容をピアフィードバックを入れたりアウトプット活動を増やしたりするなど少し変更した。さらに4年①と4年②では7月頃にコミュニケーション英語IからIIに教科書が変わり、指導方法や活動内容をAIフィードバックを入れ



■図4：模擬試験の英語全体の点数の度数分布

※Y軸の目盛りが異なることに注意



■図5：模擬試験のライティングの点数の度数分布

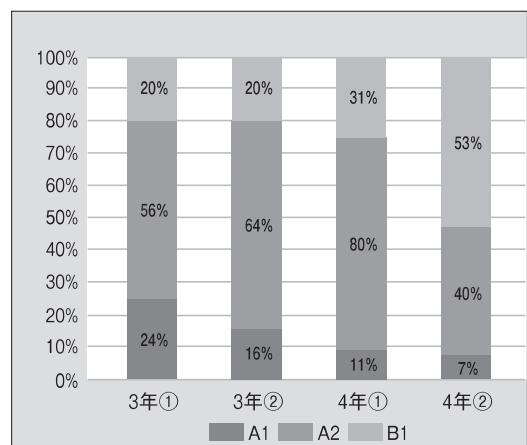
■表5：民間英語検定試験のライティングのCEFR別人数

	3年①	3年②	4年①	4年②
A1	11	7	5	3
A2	25	29	26	18
B1	9	9	14	24

■表6：表5の残差分析

	3年①	3年②	4年①	4年②
A1	2,203*	0.245	-0.734	-1.714
A2	0.173	1.555	0.518	-2.247*
B1	-1.859	-1.859	0.000	3.718**

※ *p < .05 **p < .01



■図6：民間英語検定試験のライティング力の推移

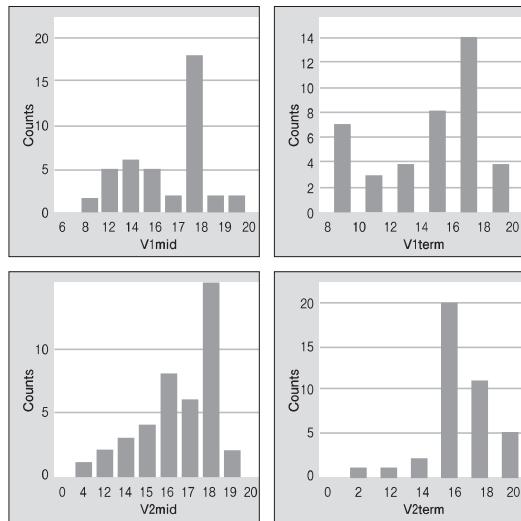
■表7：評価タスクのクラス毎の記述統計

	1中間(V1mid)		1期末(V1term)		2中間(V2mid)		2期末(V2term)	
	発展	基礎	発展	基礎	発展	基礎	発展	基礎
人数	40	14	40	14	40	14	40	14
満点	20	20	20	20	20	20	20	20
平均	16.50	13.57	15.10	10.00	16.25	15.14	16.50	16.00
標準偏差	2.36	4.09	3.74	7.72	2.63	6.69	2.93	5.14
最小値	12	6	8	0	4	0	2	0
最大値	20	18	20	18	19	20	20	20

たり、ICT教材を取り入れたり大きく変更した。4年②のB1の増加とA2の減少は指導の効果や活動の変更の影響があったと考えられる。CEFR AからBレベルへの指導の転換をねらった本調査期間の指導が効果的だと結論づけられる。

4.2.3 内部指標(評価タスク=定期考查)が示す対象者のライティング力の変化

表7は発展クラスと基礎クラスの評価タスク(定期考查のライティング自由記述問題)の記述統計である。2学期末では両クラスとも平均で16.50と



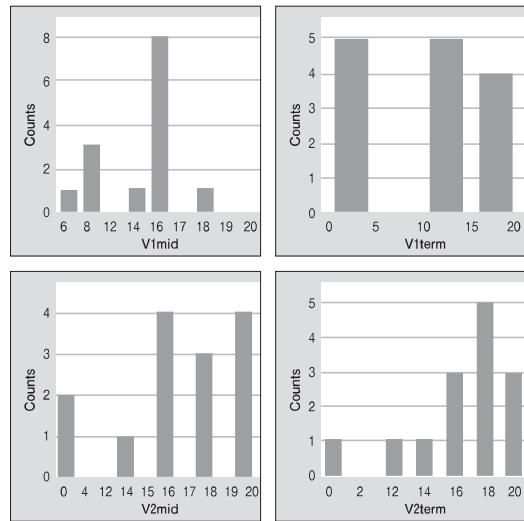
■図7：発展クラスの評価タスクの度数分布

※Y軸の目盛りが異なることに注意

16.00のように8割以上の得点率を示していることから、学習したことが定期考査で発揮できたことが分かる。一方で、図7・8を比較すると、各考査から読み取れる情報が特に基礎クラスで不安定である。例えば、1学期末で5点未満の者が他の考査より多いが、テストの難易度の影響より、試験科目に暗記教科が重なった日における実施の場合など、基礎クラスでは生徒が考査にかける学習時間の影響を受けやすく、対策の有無が大きく反映されたためと考えられる。指導と評価の一貫性やRubricの信頼性、妥当性を今後も検討する必要がある。

4.2.4 外部指標と内部指標の関連

表8は評価タスクをW&Iを使用してCEFR判定した41名分の結果を示したものである。「2期末R」は2学期末の評価タスクをW&Iを使って校正したもので、教師の直接的フィードバックなしでもスコアを返却するだけで生徒自身でW&Iを使って校正すると、A2レベルは15人が5人と減



■図8：基礎クラスの評価タスクの度数分布

り、B1レベルは20人が28人、B2レベルは3人が6人と増え、A2レベル以上の生徒はワンランク上に達することができることがわかる。表9は外部指標（リガスキル）と内部指標（12月と2月の定期考査）のCEFR判定の変化（ダウン（↓）、変わらない（→）、アップ（↑））を示したものである。GTECは7月、2学期中間とリガスキルは12月、2学期末は2月受験である。A1 = 1, A2 = 2, B1 = 3, B2 = 4に換算し、外部指標と内部指標の相関を分析した結果、2学期中間とリガスキル（表9、②）、リガスキルと2学期末（表9、③）に比較的高い正の相関が見られた（② $r=.474$, ③ $r=.402$ ）。②と③で相関が見られ、ばらつきが少ないとから、CEFR-Jを使った定期考査のライティング問題と民間英語検定試験がある程度関係付けられ、W&Iを使った日常的タスクが効果的であったことが分った。また、表8で1学期末ではB2は0名だが、2学期中間（12月）でB2が3名、2学期末（2月）でB2が3名出ていることからも、本研究の実践を継続すると、次回の民間英語検定試験ではライ

■表8: 評価タスクのCEFR別人数表

	1期末	2中間	2期末	2期末R
A1	4	1	3	2
A2	21	18	15	5
B1	16	19	20	28
B2	0	3	3	6

■表9: 外部指標と内部指標のCEFR判定の変化

GTEC ① 2中間 ② リガ ③ 2期末	↓	5	9	10
→	19	24	22	
↑	17	8	9	

■表10: 練習タスクのクラス毎のCEFR判定の変化(N=54)

		Lesson 1	Lesson 2	Lesson 3	Lesson 4	Lesson 5
発展 (40人)	↓	0	0	0	0	1
	→	28	27	23	26	27
	↑	5	9	15	7	10
	未	7	4	2	7	2
基礎 (14人)	↓	0	0	1	0	1
	→	5	9	7	8	8
	↑	0	0	1	3	0
	未	9	5	5	3	5

ティングにおいてB2の生徒が出ることが予想できる。

4.2.5 内部指標(練習タスク)が示す対象者のICTスキルとライティング力の変化

表10は発展・基礎クラス毎の練習タスク5回分において、W&Iを使ったCEFR判定の変化(校正前後でダウン(↓), 変わらない(→), アップ(↑), 「未」(未完了))を示したものである。Lesson 1では図1で指示した形式で提出できなかった生

徒(未完了者)が16人(約30%)だったのに対し、Lesson 5では7人(約13%)であった。また、1校目完了までの時間はLesson 1では平均32分(時間を記入した10人の平均)だったのに対し、Lesson 5では約27分(同生徒)だった。3か月で生徒がサマリーやタイピングに慣れてきたことが伺える(Lesson 3,4は5分短縮授業のため比較不可)。

■表11: 民間英語検定試験4年①(7月)のCEFRレベル別グループにおける使用教科書毎の伸びの変化(N=45)

	A1(5人)			A2(26人)			B1(14人)		
	中学	高1	高2	中学	高1	高2	中学	高1	高2
↓	0	1	0	5	4	2	2	0	4
→	4	4	1	15	20	9	9	5	10
↑	1	0	4	6	2	15	3	9	0

4.2.6 CEFR レベル群が示すライティング力の伸びの変化

表11は民間英語検定試験4年①におけるCEFR別グループ(A1, A2, B1)において、その前後の民間英語検定試験(3年①, 3年②, 4年②)全てを受験した45名についてのCEFR判定の伸びを示したものである。3年①と3年②の間使用した教科書は『NEW HORIZON 3』(東京書籍), 3年②と4年①の間は『PRO-VISION English Communication I』(桐原書店), 4年①と4年②は『PRO-VISION English Communication II』(桐原書店)なので、それぞれ「中学」「高1」「高2」

と表記する。中学と高校の教科書の大きな違いは英文の量と語彙の難化、構文の複雑化である。教科書が変わることで変化するのは、1コマで触れる英文の量(リスニング・リーディング)とアウトプットに求められる質と量(ライティングとスピーキング)であった。1単元4~5コマの内、中学の教科書ではインプット対アウトプットの比率が3:1に対し、高校の教科書ではICTも活用することで1:1で行った。インプットは各自のペースや技能に合うようICTを活用し、アウトプットに関しては、「高1」の教科書では、発展クラスはpeer feedback活動を、基礎クラスは教師

feedback活動を基本とした。「高2」の教科書は本研究の手順で示した通りで、AI feedback活動を取り入れた。教科書毎の伸びなし群(↓→)、伸びあり群(↑)について、js-STARを使って直接確率計算を行った結果、A2の「高1」と「高2」の間(両側検定、 $p < .00$)、B1の「高1」と「高2」の間(両側検定、 $p < .00$)で、伸びなし群と伸びあり群に有意差が見られた。A2の生徒においては、「高2」の教科書を使用した活動で、伸びなし群の減少と伸びあり群の増加が人数の間で有意だったことから、AI feedbackが有効であったことが伺える。B1の生徒においては、「高1」の教科書で、伸びなし群の減少と伸びあり群の増加が有意だったことから、高1時点でのpeer feedbackの方が高2時点でのAI feedbackよりも有効であったことが伺える。A1の生徒はどの教科書を使った活動期間においても有意な伸びは見られなかった。以上の結果から、仮説2「A2の生徒はAIによるフィードバックでB1にレベルアップできる」傾向があることが結論づけられる。

4.2.7 外部指標で顕著な伸びが見られた生徒と全く伸びが見られなかつた生徒の特性

最後に、中3～高1にかけて4年①を基点に3回の民間英語検定試験(3年①、3年②、4年②)を通して顕著な伸びが見られた生徒と全く伸びが見られなかつた生徒の事後アンケートの結果を分析する。事後アンケートの全体での結果は表12の通りである。この他、「ライティング力／文法力／語彙力を高めるのに効果的だと感じたアプリや学習方法」、「ライティング活動における先生のフィードバックとAIのフィードバックを比較して感じることや思うこと」について自由記述回答を回収した。

A1の生徒は絶対数が5人と少ないが、4回すべての英語民間試験において伸びを見せなかつた生徒が1名(S1)、4年①と4年②の間にA1からB1に伸びた生徒が1名(S2)であった。

A2の生徒のうち、3年①でA1のうち4年②までにB1に伸びた生徒が2名(S3、S4)、3年①でB1だったのち4年②までにA2に落ちたものが2名(S5、S6)であった。

B1の生徒のうち、3年①でB1のうちずっとB1に留まつたものが1名(S7)であった。伸びを見せた生徒はいなかつた。

■表12: 事後アンケートの結果

	質問項目	回答数	はい	いいえ	制限有り
1-1	家には自分が使えるPCやタブレット、スマホなどのデバイスがある	52	50 (96)		
1-2	家ではインターネットが使用できる	52	49 (94)	1 (1)	2 (3)
1-3	家でWord、Excel、PowerPointやそれ相応のソフトを使用できる	52	25 (48)	19 (36)	8 (15)
2-1	家でWrite & Improveを使用する	51	24 (47)	27 (52)	
2-2	家でサプリEnglishを使用する	51	39 (76)	12 (23)	
2-3	家でサプリを使用して英文法を学習する	52	23 (44)	29 (55)	
3-1	英語学習は楽しい	50	30 (60)	20 (40)	
3-2	英語力を高めることに意欲的だ	52	44 (84)	8 (15)	
3-3	学校外で英語を習慣的に学習する	52	21 (40)	31 (59)	

※()は%

多くの生徒が各試験でレベルを上げたり下げたりする中で、前記の7名は特筆すべきだと考え、事後アンケートの回答やそれに関連する情報を述べる。表13はこの7名の生徒の民間英語検定試験と模試のライティングのスコア、および評価

タスクのライティングのスコア、練習タスクのCEFR判定(校正後)を示したものである。練習タスクの「0」はタスク未完了を、「1」は不備がありCEFR判定は出なかつたがタスク完了を示す。

■表13:民間英語検定試験4回において顕著な伸びを示した生徒と伸びなかつた生徒のライティング力の変化

生徒	クラス	民間英語検定試験				模試				評価タスク				練習タスク				
		3年①	3年②	4年①	4年②	K5W	K8W	B7W	B11W	1中間	1期末	2中間	2期末	L1	L2	L3	L4	L5
満点				30	30	20	20	20	20	20	20	20	20					
S1	基礎	A1	A1	A1	A1	0	0	4	0	6	2	0	12	0	1	0	A2	0
S2	基礎	A1	A1	A1	B1	1	3	4	1	16	18	20	20	1	B1	B1	B1	B2
S3	基礎	A1	A2	A2	B1	10	0	8	2	14	0	16	18	0	1	B1	B1	0
S4	発展	A1	A2	A2	B1	16	13	9	14	14	18	18	16	B2	C1	B1	B2	B2
S5	発展	B1	A2	A2	A2	16	14	14	8	17	14	15	20	B1	B1	B1	B1	0
S6	発展	B1	A2	A2	A2	17	14	13	9	18	18	18	16	B2	B1	A2	B1	B2
S7	発展	B1	B1	B1	B1	27	30	14	10	18	19	18	16	B2	C1	B2	B1	B2

注:S1, S3, S5, S7は資料2参照

A1レベルで4回の試験を通して伸びなかつたS1はライティング力を高めるのに『ジーニアス総合英語』が有効だと回答したが、朝テストではほぼ毎回不合格で、これを使った学習を十分しているように思われる。また、「AI feedbackは質問できないから先生のfeedbackがよい」と回答した。A1レベルでB1まで伸びたS2はライティング力を高めるのにWrite & Improveが有効で、AI feedbackの方が結果が良い(思ったより良い結果が出る)と回答した。A2レベルでB1に伸びたS3はライティング力を高めるのに先生の丁寧な文法説明とAI feedbackが有効だと回答した。この3名が基礎クラスに所属し、模試では高い点数を取れなかつたが、評価タスクの結果からS2, S3は教師の助けを借りながら定期考查対策を割としっかりとしていたようだ。また、S2は練習タスクにもしっかり取り組んでいたことが分かる。これらの結果から、基礎レベルでは模試の点数には反映されないが、教科書付属のモデル英文を使ってAI feedbackをもらいながらサマリーを書く練習をし、Bレベルまで校正することができると、いずれは外部試験でもB1を取れる可能性が高まると想定できる。英文は文法的な誤りが多く、基本的な英文法が身についているとは断言できないが、短時間である程度の量の英文を書くことに慣れることとAIへの信頼感と教師による説明がCEFR AからBレベルへ上げる秘訣かもしれない。

S4, S5, S6, S7はライティング力を高めるの

にW&Iが有効だと回答したが、feedbackに関してはS4は「早くて正確」、S7は「早さ」という理由でAIを支持し、S5は「分かりづらい」、S6は「先生がよい」という理由でAIを支持していない。A1からB1に順に伸びが見られたS4と、ずっとB1を維持したS7は、練習タスクにおいてはAI feedbackを活用してCレベルまで校正することができた場合があったようだ。一方、3年①でB1を出してから2年経過してもA2に下がったままのS5、S6は、練習タスクにおいて課題が未完了だったり、AI feedbackで安定してB2レベルを出せなかつたりしたようだ。教師の明示的説明への信頼(AIへの不信感)を示している点もこの2名に共通している。

民間英語検定試験4年②ではB2の生徒が出なかつたが、4.2.4にも示した通り、次の試験ではそれが出ると予想できる。AI feedbackによりそれが可能となるのか今回の研究では示されなかつたが、AI feedbackの良さは「早さ」と「正確さ」である。ライティング指導をためらう理由が「添削の大変さ」と「基準の不明確さ」であるなら、AI feedbackを活用することを検討する価値があると思われる。ただし、B1の生徒をB2にレベルアップさせるには、peer feedbackや教師feedbackも取り入れながら、複雑な文法をえて多く使用するような場面や活動を設定する必要があると考える。

5

結論と今後の課題

5.1 AI feedbackの効果

本研究ではライティング力をCEFR AからBレベルへ上げることを狙った指導の転換に向けて、AIを活用した指導の効果を検討した。民間英語検定試験(3年①, 3年②, 4年①, 4年②)のライティングの伸びについてA1, A2, B1の人数の比較をすると、4年②のB1の人数の増加が有意に多かった(4.2.2参照)。また、指導方法の違いによるライティングの伸びを4年①のA1, A2, B1群ごとに人数比較すると、A2群において本研究実施期間で伸びが見られた生徒数が有意に多かった。以上のことから仮説2「A2の生徒はAIによるフィードバックでB1にレベルアップできる」可能性が高いことが分かった(4.2.6参照)。また、A1, B1群においては有意な伸びが見られなかったことから仮説1と仮説2「A1とB1の生徒はAIによるフィードバックでレベルアップできる」可能性は少ないことが分かった。

A1群でAI feedbackの効果が見られなかつた一つの原因として、文法力の欠如が考えられる。AIは直すべき候補を提示するが、どう直せばよいのかについての知識が不十分なため、校正した英文に非文が多い(資料2のS1参照)。また、練習タスクを分析すると、W&Iを使ってAI feedbackをもらってもBレベル以上になる英文への校正ができる場合が多いことが分かった。教科書本文も付属のモデル文もBレベル以上だが、その語彙を使っても練習タスクでBレベルを出せないなら、文の産出に関わる文法知識を明示的かつメタ認知的に教師がフィードバックする方が効果的だと考えられる。A1群でAI feedbackの効果が見られなかつた別の原因として、対象者数の少なさも考えられる。対象者数を増やして中期的に研究を継続すれば、別の結果が得られるかもしれない。これはB1群にも当てはまる。

A2群が短期間でB1にレベルアップできたのは、教科書がB2やC1レベルなので、教科書が後押ししたことが考えられる。AIの助けを借りながら教科書のサマリー作成を大量に行つたことが効

果的だったのだろう。また、本校の文法指導について補足したい。対象者には文法指導を授業中にはほとんど実施しなかったが、本校では前期課程(中学生段階)で高校の文法項目も先取りして、教科書『New Horizon』と文法副教材『5 Stage 英文法完成 Book 1~3』を1:1の割合で扱っている。この文法副教材はパターンプラクティスが充実しており、生徒は大量に繰り返し基本的な英文を書く習慣を前期課程で身に付ける。それに基づき、多くの生徒が後期課程(高校生段階)で文法項目を帰納的に学習したり、体系化したり、AI feedbackを読み取るための素地を作ると考えられる。この素地なくしてA2群の短期間でのレベルアップは難しかったのではないかと考える。

B1群がB2へのレベルアップをねらうとき、サマリーを量産するよりじっくり深く掘り下げた学習の方が効果的なのかもしれない。B1群においては、高1の教科書を使用してpeer feedbackの活動を取り入れた指導を行っていた頃のライティングの伸びが有意だったことがそのことを裏付けている。B2にレベルアップしたい生徒へは、表現したいことを妥協せず、B2以上の「味のある」表現に挑戦したくなるよう、生徒同士や教師のfeedbackをバランスよく取り入れ、AI feedbackに物足りなさを感じないよう工夫したい。

W&Iを使うことで、教師が見過ごしやすいスペルミスなどは生徒自ら修正でき、提出されたライティング作品はミスが減り、読みやすくなった。W&Iの瞬時のフィードバックに刺激を受けてか、ライティング活動にのめり込むように取り組んでいた生徒が多く見られたことが印象的だった。200語程度のサマリーを毎週100人単位でずっと添削するとなると気が滅入り、多分続かず、続いたとしても返却が遅れ遅れになるだろう。約2週間で1単元の4技能活動をバランスよく取り入れ、練習タスク回数を確保できたのはAIのおかげである。また、興味深いことに、日本人教師が思っている以上に文法ミスをAIは許容する。つまり、ミスの数に比例した減点をしない。A1からB1に劇的にライティング力を向上させた生徒S2が述べたように、「結果が(思ったより)良い」のである。これに励まされた生徒も多かったように思う。

本研究では、外部指標としてGTECとリンガスキルを使用したが、異なる試験のCEFR判定

の信頼性が高いという前提で分析を行った。練習タスクではW & Iを使用し、4年②で受験した民間英語検定試験はリンガスキルで、両方ともケンブリッジ大学監修である。GTECでは異なる結果になった可能性もある。また、AI feedbackなし群、peer feedback群、教師フィードバック群などの違いも本研究では扱わなかったが、今後追求したい。

5.2 ICT環境の効果

日本の高校生は「話す」「書く」のアウトプット技能の育成に課題があることが指摘されている。本研究は、多くの時間を割いてきた文法学習や「読む」「聞く」のインプット活動を見直すべききっかけとなった。ICT環境や教材を活用すれば、これらは驚くほど生徒自身でもできることができた。『EnglishCentral』や『スタディサプリ English』のおかげで「読む」「聞く」といったインプット活動時間大幅に短縮できた。教科書1単元につき、コマで読み取りや聞き取り活動を行い、1コマで読み取ったり聞き取ったりした内容からサマリーを作成し、1コマで仲間やALTとそれについてやりとりを楽しむ、という流れを作れたのはこれらのAI掲載アプリのおかげである。『EnglishCentral』は理解できるまで何度も好きなだけ視聴できることが強みであった。『スタディサプリ English』は自分のレベルに合ったものを選べることが一番の利点で、これを導入してから英語での授業に抵抗を示す生徒が減ったように感じた。そして『ジーニアス総合英語』や『スタディサプリ』の動画文法説明のおかげで、授業での文法説明を大幅にカットすることに踏み込めた。今後は、より生徒同士や教師との「やりとり」を楽しみ、「話す」「書く」のアウトプット技能も育成することが可能になるであろう。これまで時間不足やアウトプット技能が未熟だという理由で敬遠されてきた難度の高い教科書を「モデル」として触れさせることでCEFR BやCレベルの育成も可能と考える。その時には読み応え、聞き応えのある教材を生徒と共に味わいたい。

5.3 教育的示唆

外部試験導入も含めた新入試対応に向けて、現場では英語4技能をバランス良く育成するための様々な懸念や思索が見られる。本研究では、4技能をバランス良く育成することをねらい、民間英語検定試験を継続的に受験してきた中で分かってきたことを、ライティングに特化してまとめた。特に、AI feedbackを取り入れることで、これまでライティングの活動や指導において懸念されてきた、評価の懸念や添削上の課題について、新たな一つの指導法を提示することができた。CEFR-Jに基づくCan-doリストを使用することで、タスク作成における段階的指導の道筋が分かる。また、評価Rubricを学習タスク段階から提示し、身につけるべき技能をメタ認知させることで、評価タスクまでのイメージを持ちやすくなる。練習タスクでAI feedbackを活用することでライティング経験を多く積むことができる。

一方で、英語4技能をバランス良く育成しようと心がけても、模試のスコアアップには必ずしもつながらないことが分かった。大学入試のための英語力の判定に外部試験の導入が検討されているが、模試やその結果は少なからず生徒の学習動機に影響を与えている。大学が求める受験生の英語力と国が求める日本人の英語力、そしてそれを測るためにものさしが複雑化して、学習者や指導者が進むべき方向を見失うような事態は避けたい。

本研究では、定期考査における英語4技能の測定についても分析した。これも生徒にとっては普段の学習習慣に大きな影響を与えている。科目を統合し、110分で実施する等、それぞれの技能の伸長がより詳細に分析できるよう工夫した。また、解答用紙返却まで短時間しかない中で、ライティングの採点に特化できるよう他技能や知識に関する問題はマークシートを利用した多肢選択式にするなどの工夫も凝らした。

本研究から導いた結果を元に、W&IなどのICTを活用しながらCEFR Bレベルの指導法をさらに研究すると同時に、他の3技能との関連も視野に入れながら、4技能統合型の指導法も検討する段階に来ていると考える。その際には、TOEFLなどの民間英語検定試験も参考にしながら、教科書やAIを使用した指導法を開拓したい。

謝 辞

本研究は公益財団法人 日本英語検定協会の研究助成により実践できました。研究に関わる書籍の購入や研究会の参加等、助成により大変充実した研究や分析を行うことができたことをここに深く感謝申し上げます。また、小泉利恵先生から助言を頂けることになりました。小泉先生からは実践前の研究手順から微に入り細に穿つご指導を賜りました。特に一校目を仕上げてから推敲に至るまでの間、大変貴重なご指摘を賜り、私の今後の研究姿勢を見直すきっかけとなりま

した。範を垂れる姿勢を貫くよう努力します。最後に、本研究の対象者として真摯に課題に取り組んでくれた生徒の皆さんに。教師がフィードバックしないということに不安を感じなかったと言えば嘘になります。本当にAIのフィードバックだけで生徒のライティング力は伸びるのだろうか…と。結果的に予想以上の成長を示してくれたことに安堵すると同時に、大量に書かれた英文を読んで本研究成果を次なる実践につなげていこうと強く決意しました。ありがとう！

参考文献(*は引用文献)

- * Brown, M. (2018) "English-learner perceptions of Automatic Writing Evaluators" 2017年度言語メディア教育センター年報, 127-137. 神田外語大学(令和2年7月24日閲覧)
https://www.researchgate.net/publication/328094187_English-learner_perceptions_of_Automatic_Writing_Evaluators
- Nishimura, K. (2013). Demotivating Factors in Writing Classes that Utilize Online Evaluation Services. *JACET Kansai Journal*, 15, 72-91.
- OECD (2018) "The Future of Education and Skills Education 2030" (令和元年10月15日閲覧)
[http://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](http://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf)
- * Ohta, R. (2008). The Impact of an Automated Evaluation System on Student-Writing Performance. *KATE Bulletin*, 22, 23-33.
- Tajino, A., Hosogoshi, K., Kawanishi, K., Hidaka, Y., Takahashi, S., & Kanamaru, T. (2011). Feedback in the academic writing classroom: Implications from classroom practice with the use of Criterion®. *Kyoto University Studies in Higher Education*, 17, 97-108.
- * Uchida, S & Negishi, M (2018) Assigning CEFR-J levels to English texts based on textual features. In Tono and H. Isahara (eds.) *Proceedings of the 4th Asia Pacific Corpus Linguistics Conference* (APCLC 2018), pp. 463-467.
- ケンブリッジ大学英語検定機構(2018)『Linguaskill ライティングトライアルレポート2017年6月』(令和2年7月24日閲覧)
<https://www.kamwai-juku.ac.jp/cambridge-english/pdf/linguaskill-writing-trial-report.pdf>
- * 斎藤雪絵 (2017)『自動採点システムを使った英語ライティング学習』立教大学ランゲージセンター紀要, 38, 63-74 (令和2年7月24日閲覧) https://rikkyo.repo.nii.ac.jp/?action=repository_uri&item_id=15456&file_id=18&file_no=1
- 平林健治(2016)『日本人上級英語学習者の自由英作文の言語的特徴の分析』中部地区英語教育学会紀要, 23-30.
- * 文部科学省(2017)『新しい指導要領の考え方』(令和元年10月15日閲覧)
http://www.mext.go.jp/a_menu/shotou/new-cs/_icsFiles/afieldfile/2017/09/28/1396716_1.pdf
- * 文部科学省(2018)『OECD Education 2030 プロジェクトについて』(令和元年10月15日閲覧)
http://www.oecd.org/education/2030/OECD-Education-2030-Position-Paper_Japanese.pdf
- * 文部科学省(2018)『英語力調査結果(高校3年生)』(令和元年10月15日閲覧) http://www.mext.go.jp/a_menu/kokusai/gaikokugo/_icsFiles/afieldfile/2018/04/06/1403470_03_1.pdf
- 文部科学省(2018)『各資格・検定試験とCEFRとの対照表』(令和2年7月24日閲覧) https://www.mext.go.jp/b_menu/houdou/30/03/_icsFiles/afieldfile/2019/01/15/1402610_1.pdf

資料1：「Can-do リスト」

新潟県立津南中等教育学校 Can-do リスト

到達時期 (めやす)	学年末の 到達目標	話すこと		書くこと	聞くこと	読むこと	英検
		やりとり	発表				
卒業後	英語を通じて社会的な幅広い話題についての情報や考え方などを的確に理解したり、適切に伝えたりすることができる。	一般的な分野から、文化、学術などの、専門的な分野まで、幅広いトピックの会話を積極的に参加し、自分の考えを正確かつ流暢に表現することができる。【B2.2】	ディベートなどで、社会問題や時事問題に関して、補助的観点や関連事例を詳細に加えながら、自分の視点を明確に展開することができ、話を続けることができる。【B2.2】	自分の専門分野や関心のある事柄であれば、複雑な内容を含む報告書や論文などを、原因や結果、家庭的な状況も考慮しつつ、明瞭かつ詳細な文章で書くことができる。【B2.2】	自然な速さで標準的な発音の英語で話されれば、現代社会や専門分野のトピックについて、話者の意図を理解することができます。【B2.2】	記事やレポートなどのやや複雑な文章を一読し、文章の重要度を判断することができる。綿密な読みが必要と判断した場合は、読み速さや読み方を変えて、正確に読むことができる。【B2.2】	Grade Pre1
	英語を通じて、言語や文化に対する理解を深め、様々な話題についての情報や考え方などを的確に理解するとともに、積極的にコミュニケーションを図ろうとする態度を身につけ、自分の考えなどを適切に伝えたりすることができる。	母語話者同士の議論に加われないことがあるが、自分が学んだトピックや自分の興味や経験の範囲内のトピックなら、抽象的なトピックであっても、議論できる。【B2.1】	ディベートなどで、そのトピックが関心のある分野のものであれば、論拠を並べ自分の主張を明確に述べることができる。【B2.1】	そのトピックについて何か自分が知っているは多く情報源から統合して情報や議論を整理しながら、それに対する自分の考えの根拠を示しつつ、ある程度の結論性のあるエッセイやレポートなどを、幅広い語彙や複雑な文構造をある程度使って、書くことができる。【B2.1】	トピックが身近であれば、長い話や複雑な議論の流れを理解できる。【B2.1】	現代の問題など一般的関心の高いトピックを扱った文章を、辞書を使わずに読み、複数の視点の相違点や共通点を比較しながら読むことができる。【B2.1】	
	ある程度なじみのあるトピックならば、新聞・インターネットで読んだり、テレビで見たニュースの要点について議論することができる。【B2.1】	自分の関心事であれば、社会の状況(ただし自分の関心事)について、自分の意見を覚えて、自分の意見をすらすらと発表し、聴衆から質問がでれば相手に理解できるように答えることができる。【B1.2】	物事の順序に従って、旅行記や自分史、身近なエピソードなどの物語文を、いくつかのパラグラフで書くことができる。また、近況を詳しく伝える個人的な手紙を書くことができる。【B1.2】	自然な速さの録音や放送(天気予想や空港のアナウンスなど)を聞いて、自分に関心のある、具体的な情報の大部分を聞き取ることができる。【B1.2】	平易な英語で書かれた長めの物語の筋を理解することができる。【B1.2】	Grade2	
	英語を通じて、言語や文化に対する理解を深め、身近にある様々な話題についての情報や考え方などを的確に理解するとともに、積極的にコミュニケーションを図ろうとする態度を身につけ、自分の考え方や他の人の考え方などを伝えたりすることができる。	個人的に関心のある具体的なトピックについて、簡単な英語を多様に用いて、社交的な会話を続けることができる。【B1.1】(3分程度)	"自分の考えを事前に準備して、メモの助けがあれば、聞き手を混乱させないように、馴染みのあるトピックや自分に関心のある事柄について語ることができ。【B1.1】(3分程度)"	自分に直接関わりのある環境(学校・職場・地域など)での出来事、身近な状況で使われる語彙・文法を用いて、ある程度まとまりのあるかたちで、描写することができる。【B1.1】(1段落)	外国の行事や習慣などに関する説明の概要を、ゆっくりはっきりと話されれば、理解することができる。【B1.1】		学習の目的として書かれた新聞や雑誌の記事の要点を理解することができる。【B1.1】
	予測できる日常的な状況(郵便局・駅・店など)ならば、さまざまな語や表現を用いてやり取りができる。【A2.2】	写真や絵、地図などの視覚的な補助を利用しながら、一連の簡単な句や文を使って、自分の毎日の生活に直接関連のあるトピック(自分、学校や地域など)について短いストーリーをすることができる。【A2.2】(2分程度)	聞いたり読んだりした内容(生活や文化の紹介などの説明や物語)であれば、基礎的な日常生活語彙や表現を用いて、感想や意見などを短く書くことができる。【A2.2】(100語程度)	学校の宿題、旅行の日程などの明確で具体的な事実を、はっきりとじみのある発音で指示されれば、要点を理解することができる。【A2.1】	簡単な語を用いて書かれた人物描写、場所の説明、日常生活や文化の紹介などの説明文を理解することができる。【A2.1】	Grade Pre2	
	初歩的な英語を通じて、言語や文化に対する理解を深め、話し手や書き手の意向などを理解するとともに、積極的にコミュニケーションを図ろうとする態度を身につけ、自分の考え方などを話したり書いたりすることができる。	趣味・部活動などのじみのあるトピックについて、はっきりと話されれば、簡単な質疑応答をすることができる。【A1.3】(1分程度)	前もって発話することを意図して、限られた身近なトピックについて、簡単な語や基礎的な句を限られた構文に用い、複数の文で意見を言うことができる。【A1.3】(1分程度)	趣味や好き嫌いについて複数の文を用いて、簡単な語や基礎的な表現を使って書くことができる。【A1.3】(60語程度)	ゆっくりとはっきり話されれば、自分自身や自分の家族・学校・地域などの身の回りの事柄に関連した句や表現を理解することができる。【A1.3】	"じみのあるトピックについての英文なら初見で20秒の默読リハーサル後に聞き手に意味が伝わるよう音読することができる。"	
1年	なじみのある定型表現を使って、時間・日にち・場所について質問したり、質問に答えたりすることができる。【A1.1】	基礎的な語句、定型表現を用いて、限られた個人情報(家族や趣味など)を伝えることができる。【A1.1】	自分について基本的な情報(名前、住所、家族などを)を辞書を使えば短い句または文で書くことができる。【A1.1】(5文程度)	日常生活に必要な重要な情報(数字、品物の値段、日付、曜日など)をゆっくりと話されれば、聞きとることができます。【A1.1】	既習の英文なら、日本語の助けを借りて、レポーターのように暗唱することができる。	既習の英文なら、日本語の助けを借りて、レポーターのように暗唱することができる。	Grade3

資料2：顕著な伸びがあった生徒と伸びがなかった生徒の練習タスクに提出された英文
※基礎クラスの1st draftは教科書付属のモデル文(Lesson 4 Handwriting in the Digital Age)

【S1：基礎クラス・民間英語検定試験は4回ともA1】

(First draft)

The author is saddened by the dwindling attention that school are giving to the skill of handwriting. Computers affect our writing. The auto-correct spelling function in our computers makes our spelling ability worse. In addition, the Internet hinders our progress of writing. There are still advantages of handwriting in our digital age today. Handwriting turns out to be especially effective when it comes to brainstorming, through which we can come up with various ideas, and connect and group them. Moreover, handwriting today has its own place. We should take advantage of handwriting or a keyboard, depending on what we write, why we write, and to whom we write/

108 words 12:58 B2

(Second draft)

the author is saddened reason to the dwindling attention that school are giving to the skill of handwriting. The auto-correct spelling function in our computers makes our spelling ability worse. The Internet mainly hinders our process of writing. When it comes to brainstorming to handwriting turn out to be especially effective. It enables one to come up with variouse ideas, and connect and group them. It enables of handwriting enable one to be. Better understanding and memorization of the class material to handwriting lead to. We need to consider what we write, why we write, and to whom we write to handwriting in the digital age.

110 words A2

【S7：発展クラス・民間英語検定試験は4回ともB1】

(First draft)

This is a story about handwriting. The author is saddened by the dwindling attention that schools are giving to the skill of handwriting, putting greater emphasis on developing keyboard skills instead because he is grateful that there was so much time and attention devoted to handwriting and penmanship and we can usually write and edit our papers with the word processing application. When we make spelling mistakes on the computer, the computer correct our mistakes automatically, therefore, our spelling ability is getting worse and worse in these days. Times that we memorize the spelling without the computer are important for us. Actually, handwriting turns out to be especially effective when it comes to brainstorming, that means, it enable us to think without any restrictions. Moreover, handwriting leads to better understanding and memorization of the class material. We need to consider what we write, why we write, whom we write in choosing handwriting or a keyboard for protecting our own ability.

160 words 20 minutes B1

(Second draft)

A man who appreciate that there was so much time and attention devoting to handwriting and penmanship was saddened by giving to the skill of keyboard instead of handwriting at school. When we make spelling mistakes on the computer, the computer correct our mistakes automatically or indicated by red underlines. By using this ability, our spelling ability is getting worse and worse in these days. We must have time to memorize word' spelling perfectly without the computer . As a matter of fact, according to the research, handwriting turned out to be especially effective when we want a new idea. That means, it enable us to think without any restrictions. Moreover, handwriting allows us to process the class material more deeply, which leads to better understanding and memorization. From this fact, we need to consider what we write, why we write, whom we write in choosing handwriting or a keyboard for protecting our own ability.

156 words B1

資料2: 顕著な伸びがあった生徒と伸びがなかった生徒の練習タスクに提出された英文

※基礎クラスの1st draftは教科書付属のモデル文(Lesson 4 Handwriting in the Digital Age)

【S3 : 基礎クラス・民間英語検定試験は A1 → A2 → A2 → B1】

(First draft)

(モデル文省略)

11:21 B2 108 words 12:58 B2

(Second draft)

The author notes the decline in handwriting skills due to the development of computers. Computers affect writing. The computer's auto-correct spelling feature reduces the spelling ability of the computer's spelling capabilities. In addition, the Internet hinders writing. In the digital age today, there are still handwriting benefits. Handwriting proves to be particularly effective for brainstorming. By thinking with your own head, you can come up with various ideas. In addition, handwriting leads to better understanding and memorization of class materials. Today's handwriting role has its own place. You need to use handwriting or keyboard, depending on what you write, why you write it, and who you write to.

108 words B1

【S5 : 発展クラス・民間英語検定試験は B1 → A2 → A2 → A2】

(First draft)

These days, people don't write with a pen. A author is saddened by dwindling attention that schools are going to the skill of handwriting.

The auto correct spelling function in our computer makes our spelling ability worse. The mainly hinders our process of writing is "Internet".

What is the good point of handwriting is it turns out to be especially effective when it comes to brainstorming. Also it enable to think without any restrictions.

Handwriting leads to understanding and memorization. Now we need to consider what we write, why we write, and to whom we write.

96 words 15minutes A2

(Second draft)

These days, people don't write with a pen. An author is saddened by dwindling attention that schools are going to the skill of handwriting.

The auto-correct spelling function in our computer makes our spelling ability worse. The main hinders our process of writing is "Internet".

What is the good point of handwriting is it turns out to be especially effective when it comes to brainstorming. It also enable to think without any restrictions.

Handwriting leads to understanding and memorization. Now we need to consider what we write, why we write, and to whom we write.

96 words B1

B 実践部門・報告Ⅳ・英語能力向上をめざす教育実践

Pre-reading でのトップダウンアプローチと Post-reading の統語処理活動を取り入れた 高校生の多読授業実践

研究者: 東京都／東京都立 国分寺高等学校 教諭 三上 洋介(申請時: 東京都立 大田桜台高等学校)
《研究助言者: 長 勝彦》

概要 本実践研究では、校種や生徒レベルを問わずに活用でき、多読活動時に生徒の興味を喚起し、総合的な英語力を向上させるためのワークシート型多読学習の支援ツール(Odyssey Reading Worksheet=ODR)の開発を試みた。このODRは、生徒の興味を喚起させる内容推測などのトップダウン処理と、サマリーライティングやペア活動などの統語処理活動を取り入れている。

ODRを用いた多読を実践したところ、処置群には外部試験のリーディングスコアの上昇や速読スピードの向上が見られた。また、多読読書語数は上昇し、ワークシート型の多読による「読みの変化」や統語処理のアウトプット活動による多読意欲の向上も示唆され、アンケート結果から情意面の変化も見られた。学習者の背景知識を活性化させることで、英文の読書負担の軽減、本文の理解度の向上や、統語処理活動による未知の単語を読み飛ばす頻度が低下するなど、英文を深く捉える傾向につながったと言える。さらに、従来の多読に比べて意欲が高まり、想像力が刺激されるなど、学習者はODRを肯定的に評価した。

今後の課題として、学習者の更なる動機付けのため、ODRの理論的背景や意義を学習者へ理解させることや、ワークシートの使用に伴って減少する授業内多読時間への対処が挙げられる。また、効果的な使用法の提案として、小中学生にも使用可能な簡易版を開発することや、多読図書に限らず、英語雑誌や評論文など多様な英文素材読書へODRを応用することも検討できると考えられる。

1 多読授業について

1.1 多読授業の効果と課題

多読は、精読と異なり、簡単な英語で書かれた本を大量に読む、楽しみのための読書活動である。学習者は自分で本を選び、1人で読書する。指針としてよく取り上げられるのが、(1)辞書は引かない(英語は英語のまま理解する。日本語に訳さない)。(2)わからないところはとばす(7~9割の理解度で読む)。(3)つまらなければあとまわし(内容自体を楽しもう)(酒井, 2002, 古川, 2010)などの「多読3原則」だ。

これまで多読の利点は数多く報告されている。精読に比べて圧倒的な英語のインプット量を確保できるので、読みの流暢性が向上する。英語への自信や読書習慣が身についたり、世界や文化の知識を獲得できるなどのメリットもある。

多読授業は日本国内で急速に普及しており(Furukawa, Takase, & Nishizawa, 2009), 学校教育での導入数も増えているが、日本での成功例はまだ少ないといわれる。その原因として(1)読書量不足(10か月間の海外留学に相当する語数は200~300万語)、(2)英文レベルの選択ミスがある(西澤・米澤・粟野, 2019)。継続すれば効果があるにも関わらず、多読への意欲を維持できず、十分な読書量を確保できない。また、自分のレベルに合った選書が出来ず、英文を理解できないので、

成果が見える前に多読をやめてしまうのである。

また、多読授業実践の成功例や実践研究対象は、ある一定の学力水準の公立・私立高校や高等専門学校、中高一貫校などが多い(高瀬, 2010)(増見, 2014)(中野, 2014)。これは多読授業が、読書習慣や基礎学力を有する学習者に対して親和性が高く導入しやすいためであろう。一方、いわゆる進路多様高校における多読の実践研究は極めて少ない。これには生徒の読書習慣確立から取り組まなければならず、多読学習の成果を上げるために多大な時間と労力が必要とされることが背景にある。これらの点から、個々の学校実態に合った多読授業の実践が課題となっている。

2

大田桜台高校の 多読多聴授業

都立大田桜台高校は東京都で2校指定されている「ビジネスコミュニケーション科」の1つである。実社会で使える英語力とビジネス知識の習得を目指し、生徒は1年生から実践的な学習に取り組んでいる。特に英語に関しては、3年間の必修科目は20単位と多い。

2.1 図書室で行う多読授業

多読授業は図書室で行われ、「図書室多読」とも呼ばれる。読書とは本来個人的な活動であるが、読書習慣のない生徒にいきなり個人で洋書を読ませることはハードルが高い。まして授業外で洋書の読書を生徒に期待することは難しい。そこで読書習慣を根付かせ、多読の楽しさを実感させるために図書室多読の手法を取っている。図書室には英語の絵本や語彙制限のある洋書など、約14,000冊が蔵書されている。50分の多読授業中、生徒は自由に本を選び、読書を進めていく。

ただし、図書室で多読授業を行っても、生徒が自発的かつ継続的に読書を行うまでにはかなりの時間を要する。「目の前に本を置いただけで本を手にして読む生徒がいたら、その生徒はよほどの本好きか、読書の習慣が形成されているといえる」(江竜、村松, 2018)というのが、多読を実践し

ている多くの教師の本音である。

2.2 CALL 教室で行う多聴授業

多読と両輪を成す多聴授業は CALL 教室で行われる。多聴とはパソコン設置の教室で、1人につけ1台のPCが用意され、朗読音声を聞きながら本を読むことである。英語の音声やリズムに合わせて読書し、英語を英語のままで理解する習慣を作る。この CALL 教室には約1,000～1,200冊の蔵書が備えられ、PCのサーバ内に蔵書分の音声データが格納されている。多読授業で読んだ本を選び、多聴授業で音声を聞きながら読むことを推奨している。

2.3 多読多聴授業のツール

生徒の多読多聴学習を支援するため、いくつかの支援ツールを活用している。

(1) 多読多聴手帳

多読多聴手帳とは、読書履歴を記録するためのもので、生徒1人につき1冊配布する。1冊読み終えるごとにタイトルや感想を多聴手帳に1行で記録する。多読時は読んだ本のワード数、多聴時は聞いた音声時間とワード数を記録し、累計読書語数や累計リスニング時間も管理できる。教師は生徒の読書履歴を確認して、激励コメントやおすすめ本の紹介などを書き込む。英語多聴へのモチベーションを保つことを目的としている。(写真1)。多読多聴手帳は、見開き2ページに20冊分を記録でき、この手帳1冊に1,000冊分の読書記録が書き込める。20冊読了するごとに、累計読書語数を集計させる。

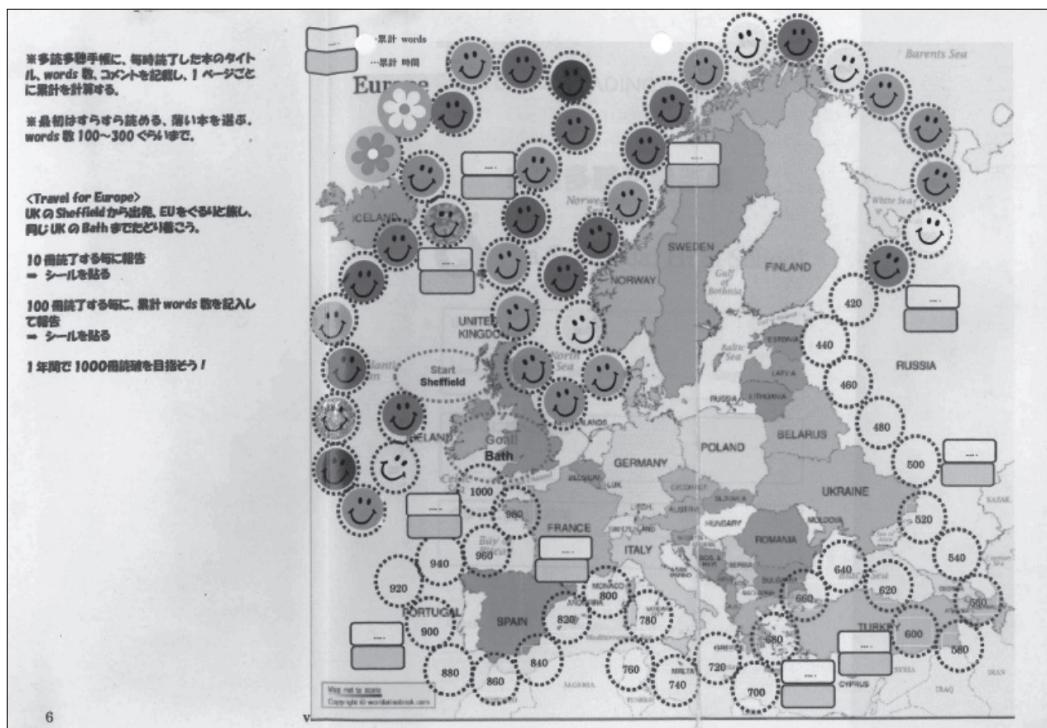
(2) 多読多聴地図

(多読多聴冊数の可視化ツール)

多読多聴地図とは、累計多読冊数を可視化し、薄い洋書を大量に読むことを促すための地図である(写真2)。10冊読了するごとにシールを配布し、地図に貼る。1,000冊読み終えると、地図が完成する。この地図は多読多聴手帳の裏表紙に貼りつけることで、多読の進捗状況を把握でき、多読の達成感を得ることが出来る。

No.	Date	Series	Title	Words	Minutes	Comment
1621	11/15	RG3-16	The Littles	6,739		2910字で読みました。
1622	11/16	RG3-15	Lily and miss Liberty	6,679		66の字で読みました。
1623	11/16	RG2-15	Cannibals don't eat me well	6,130		6,130字で読みました。
1624	11/14		Camel Back mountain	358		山を読みました。
1625	11/14	SIRY-6	The titanic lost and found	2,094		映画を読みました。
1626	11/17	WFH1-07	It can't be	1,021		姉妹が書いた物語です。
1627	11/17	WFH1-08	People like that	1,140		Najima was asked for a
1628	11/17	WFH1-09	I hate computers	1,092		PCで書かれた物語です。
1629	11/17	WFH1-10	Toxic waste	1,082		児童小説です。Tori
1630	11/17	SIRY-53	The great fairy race	660		All the fairies run to the
1631	11/17	SIRY-1	Helen Keller can't see	1,950		Helen Keller's story
1632	11/17	AAR2-22	Sara jawea her twin	1,300		Sara jawea's story
1633	11/18	ORTM-12	The joke machine	888		Children reading
1634	11/18	ORTM-13	Sjorarine Adventure	890		wife a pre-adolescent
1635	11/18	CJ-11	Com Jason the mystery of the stolen orange	5,376		1-12歳で読みました。
1636	11/18	CJ-12A	Com Jason the mystery of light 54	5,210		54種類に分けてあります。
1637	11/18	CJ-13	Com Jason the secret of the haunted house	5,429		家が怖いとか、怖がる。
1638	11/19	CJ-14a	Com Jason the chocolate Eudora mystery	5,523		1-12歳で読みました。
1639	11/19	CJ-15	Com Jason the secret of pale knight	5,045		5,045字で読みました。
1640	11/19	CJ-16	Com Jason the giant y mystery	5,284		外国の本を読みました。
MEMO				64,503	0	←このページの合計
				920 / 62	3,153	←今までの合計

■写真1：多読多聴手帳の記入例



■写真2：多読多聴地図

(3) 多読多聴学習の導入用必読書リスト

英語に苦手意識を持つ読書習慣も持たない生徒が、スムーズに英語多読を開始できるよう、多読導入用必読書リストを準備している。使用する絵本は、英国の多くの小学校の国語授業で使用される Oxford Reading Tree シリーズである。この Stage1-3 の 96 冊を全生徒は高校入学時から 3か月程度で読み終える。また、多聴授業では、同じ Oxford Reading Tree の音声 CD を聞きながら多聴する。これにより英語多読に慣れ、英語を英語のままで理解する土台作りを目指している。

(4) 推奨読書リスト

導入用必読書リストを読み終えた生徒は、推奨読書リスト(資料1)から好きな本を選んで読書をする。難易度の高すぎる本を選ぶなど選書を誤ると、多読学習の効果も上がらないため、こうした推奨読書リストを作成している。

推奨読書リストの核は Oxford Reading Tree である。多くの生徒は Stage1-4 までは読了できるが、Stage5 以降は難易度が高いと感じる生徒が急激に増える。そのため Stage4 付近で、易しくて薄い洋書を大量に読むことを推奨している。徐々に難易度を上げ、最終的には Stage9 まで読了することを促す。

2.4 指導方法

2.4.1. 多読導入時のオリエンテーション

英語に苦手意識を持つ生徒が大半を占める当校では、スムーズに多読学習に入れるよう入学時に「多読オリエンテーション」を実施している。毎年 NPO 多言語多読の酒井邦秀氏を招き、新入生全員に Oxford Reading Tree の Stage2 または 3 が 1 冊配布され、多読の方法を体験させる。まず、タイトルや表紙の絵をじっくりと眺めさせる。それから挿絵を頼りに、日本語に訳さず英語を英語のままで読むよう指導する。わからない単語にこだわりすぎず、絵から推測することを推奨し、まずは内容を楽しむことを重視させている。このオリエンテーションで、辞書を使わずに洋書を読み進めることも覚え、生徒は多読の本の読み方を身につける。

2.4.2. 多読の指導

多読指導の方針は多読 3 原則に沿って、できるだけ易しいレベルの絵本から読み始め、少しづつレベルアップを体感させることを最重視している。生徒には「できるだけ薄く、スラスラ読める洋書を読もう。まずは 500 冊、1000 冊を目指そう」と、累計語数よりも冊数重視の指導をしている。たとえ薄くとも洋書 1 冊を読み切ることで、達成感を得やすく、やさしい英語を早く読む訓練にもなる。中学英語の基礎が未定着の生徒には、このような冊数重視の目標設定を行う。

生徒に読書習慣を身につけさせるため、「読書中は静寂を作る」「姿勢を正して読む」「多読多聴手帳に記録する」といった多読授業を受けるルールも別途設定している。通常の授業と異なる、集中して多読する「場づくり」を行うためである。

2.4.3. 授業の流れ

標準的な 50 分の多読授業の流れは、冒頭の 3~5 分で生徒自身が選書し、その後は読書する。当校の多読多聴授業には決まった活動はない。教師側の役割は、読書に集中できる「場づくり」と、生徒個人への助言である。多読多聴手帳を読んで生徒の読書傾向を把握し、適当な本を推奨する程度である。多読授業に変化を持たせたい場合、目の前の生徒実態に合わせ、教師が必要な活動をその都度取り入れる。例えば、ブックトーク(本の紹介活動)や本の読み聞かせを取り入れることもある。

2.5 英語授業のカリキュラム上の特徴

大田桜台高校では、「使える英語を楽しく身に付けさせる」ことを目指し、全学年カリキュラムとして必修授業に多読・多聴学習を導入している(授業内多読授業)。平成 21 年の開学以来、全英語授業(コミュニケーション英語や英語表現科目)コマ数の半分が多読・多聴授業になっており、残りの半分は教科書を使った教室授業である。1 年次は 7 単位(コミュニケーション英語 I : 4 単位、英語表現 I : 3 単位)、2 年次は 8 単位(コミュニケーション英語 II : 4 単位、英語表現 II : 4 単位)、3 年次は 5 単位(コミュニケーション英語 III : 5 単位)である。

具体的には、1年次は多読授業:2単位、多聴授業:2単位、2年次は多読授業:2単位、多聴授業:2単位、3年次は多読授業:1単位、多聴授業:1単位が設定されている。つまり3年間で多読授業は5

単位、多聴授業も5単位受講する計算である。これにより生徒は大量の英語に触れる機会を確保できる(表1)。

■表1: 2019年度 1週間の英語必修科目的授業コマ数(1コマ50分)

	1	2	3	4	5	6	7	8	授業数計
1年次	コミュニケーション英語I				英語表現I				7
	ER	EL	J	J	ER	EL	J		
2年次	コミュニケーション英語II				英語表現II				8
	ER	EL	J	J	ER	EL	J	J	
3年次	コミュニケーション英語III								5
	ER	EL	J	J	J				

ER=多読、EL=多聴、J=日本人教師授業

2.6 多読多聴学習の効果測定

多読多聴学習の効果測定にGTEC for STUDENTSを使用している。受検時期は高校1年次の4月、高校2年次の7月、高校3年次の7月、12月の4回である。学年平均のGTEC(3技能)のスコアを、高校入学時から卒業までに100点以上向上させることを目標としている。

2.7 多読授業の課題

本校での多読授業実践も4年目を迎え、生徒の英語力向上に手ごたえを感じる一方、次の3つの問題意識を持つに至った。第1に、いかに生徒の読書意欲を喚起するかである。多読学習に親和性を示す生徒がいる一方で、読書意欲を持てない生徒も一定数存在する。多読授業の敵は、「マンネリ」や「飽き」であり、いったん生徒が多読学習に飽きれば、個人的な読書活動である多読授業はその意義を失う。特に、AO入試等で早期に進路決定する高校3年生の後半は、生徒の多読学習意欲が大幅に低下する傾向がある。そこで洋書の読書意欲を喚起する仕掛けや、小さな達成感を実感できる多読授業の展開が不可欠である。

第2に、多読を重ねるもの英語力の伸長が鈍い学習者の存在である。例えば、単に文字や絵をぼんやり眺め、知っている単語をつなげて内容をあいまいに理解する、統語処理をしていない読み方をする生徒も見られる。多読を10万語程読み

続けたにもかかわらず、GTECのリーディングスコアに伸びの見られない生徒は、英文を統語処理せず、なんとなく読む傾向があった。こうした理解度の低い状態で多読活動を続けていると、読書活動のインプットが英語力向上に結びつかない。英文をより深く捉えさせ、内容を深く味わうための工夫が必要であった。

第3に、多読授業の「変数」をどう抑えるかである。多読指導方法には定型が存在しない。多読は生徒の個人的な読書活動のためである。特に、多様な背景を抱える生徒の多い公立高校では、多読への取り組み姿勢は生徒によってばらつきがあり、「個別指導」的に生徒一人ひとりの様子を観察して支援する必要がある。そのため教員の定期異動がある公立高校にとって、経験を持たない教員は多読指導に戸惑うことも多く、多読指導ノウハウの標準化、蓄積や共有も常に課題となっていた。もし校種や学力水準を問わずに機能する「標準的な多読活動」があれば、教師側にとって多読指導は容易になるであろう。

3 先行研究

3.1 言語処理のプロセス

「読み」とは、文字のテキストを読むことで、書き手の意図する意味内容を理解する行為である。

そのためには、読み手は自分の持つ言語知識だけでなく、常識や背景知識といった既存知識を使いながら、「積極的にテキストに働きかけて、意味内容を構築する行動者」(天満, 1989)でなければならない。

読み手の持つ既存知識を意味する用語は、「背景知識」や「スキーマ」などがある。単にテキストの解読と考えられていた読解を、読み手がテキストを再構成する過程と捉えたのは Bartlett(1932)であり、過去に経験した蓄積を抽出した知識構造を「スキーマ」と定義した。また、長期記憶内に貯えられている知識を「スキーマ」(Rumelhart and Ortony, 1977)と位置づけることもある。

3.2 リーディングの処理過程

背景知識であるスキーマを基盤にして、リーディングの処理過程は次の2つが存在するとされている。1つはボトムアップ処理(bottom-up processing)と呼ばれ、アルファベットの文字認識から単語認識へ移行し、それらの認識が集まって大きな文や文章の理解へと段階的に処理が進行する、「細部から全体へ」という考え方である(Gough, 1972)。

もう一方はトップダウン処理(top-down processing)である。自分の頭の中にある背景知識を使い、書かれているものについて仮説を立て、読みながら自分の予測を確認する作業を繰り返す。「全体から細部へ」という考え方である(Goodman, 1967) (Rumelhart and Ortony, 1977)。つまり、文章を効率的に理解するためには、文章と自分の背景知識を関連づける能力が必要になる(Carrel and Eisterhold, 1983)。

実際の読解の過程では、両者が用いられて、優れた読み手は同時並行でボトムアップ処理とトップダウン処理を行い、内容理解に達すると考えられている。これが相互作用モデル(interactive model)である(Rumelhart et al, 1986)。

このように「読み」はボトムアップ処理とトップダウン処理の相互作用と考えられるが、津田塾(2002)は「読みの苦手な人は往々にしてボトムアップの過程だけにとどまることが多いため、全体理解に及ばないことがある」と指摘している。英語の苦手な生徒への多読指導において、トップ

ダウン処理を促すことは検討されるべきであろう。

3.3 プレリーディングの活動

スキーマの考え方を読解授業に応用し、プレリーディング時に既存の背景知識を活性化する活動を提倡したのは Carrell(1984)である。文章を読む前にトピックに関する背景知識や視覚素材を提示することやタイトルから内容を推測させる活動は広く普及し、現在でも教室のリーディング授業等で使用されている。

代表的な活動にセマンティック・マッピングがあげられる。これはリーディングを行う前に、そのテキストのテーマに関する情報を視覚的にまとめる活動である。もともとは語彙増強に主眼が置かれ、日本の中学校でも語彙指導上の効果が報告されている(北原, 2010)。近年は内容理解を深化させる点にも注目が集まり、読解前活動として学習者の背景知識を活性化し、テーマに関する自分の知識を整理し、読み手としての関心を高める利点が報告されている(卯城, 1993)。具体的には、読書前にトピックに関する語句を出来るだけ考えるよう生徒を励まし、その予測した語句を紙に書かせる。

こうしたプレリーディング活動を行う理由を、Aebersold and Field(1997)は次のように述べている。第1に、文章の理解度を深める。読書前に個人的な知識や経験、背景知識といったスキーマを活性化させると、学習者はこの知識を心に留めながら読むことになり、文章の中で見出した情報をより理解することができる機会を増やす。第2に、読む動機づけができる。背景知識を活性化し内容を推測すると、そのトピックに関する関心は必然的に高まる。第3に、文章を読む目的を確立できる。文章の内容を事前に予測することで、より効率的な読みを行うことが出来る。このプレリーディング活動において、内容を1, 2行の英文で予測するなどのライティングが取り入れされることもある。

学習者がスキーマを活性化させて内容を予測すると、筆者の見解と自分の見解がどれほど近いか、または離れているか、という視点で文章を読むことができる。これが文章の深い理解へとつながる(Zamel, 1992)。このようにプレリー

ディング活動は、学習者の頭の中に蓄積された知識を活発に働きかせる、トップダウン処理を促すといえる。

3.4 トップダウン処理の有用性

トップダウン処理を行わないと、文章を完全に理解できないという実験例がある。

実験は2グループに分けて行われ、前者は事前に主題についての情報を与えられ、後者は何の情報も与えられずに、文章を録音したテープを聞いて、内容確認や再話のタスクが与えられた。結果は前者のほうが理解度は高かった(Bransford & Johnson, 1972)。

主題の情報なしで読んだ場合、何の話かを理解することは困難である。文章の冒頭に主題が書かれていれば、既存の背景知識を使いながらテキストを読めるので、理解は容易になるであろう。文章理解における背景知識を活かしたトップダウン処理の有効性を示す一例である。この点から、多読においても、タイトルや挿絵から内容を推測することで、読者はより容易に深い読みに至るといえよう。

3.5 ポストリーディングの統語処理活動

ポストリーディング活動でよく行われるのが、内容を振り返るための、要約や再話といった統語処理を必要とする活動である。要約は、削除、一般化、選択、構成の4種類で構成される(Kintsch and Dijk, 1978)。これを行うためには、文章中の重要な情報とそうでない情報を分ける能力が必要である。このため文章への積極的な関与が読み手に求められ、結果として内容の理解を深める。しかし、ポストリーディング活動で広く用いられる要約は、第一言語での研究が中心であり、特に、第二言語における研究が極めて少ない(門田、野呂, 2001)といわれる。

また、再話もポストリーディング活動によく用いられる。読書後に、本文を見ずに、相手に対して内容を語る活動である。再話する側は、自分の頭の中で英文内容を整理し、再構築しなければならない。「この過程が、テキストから得た情報と読み手の背景知識を統合する読解プロセスと類

似している」(卯城, 2009)ため、再話を通じて読み手は筋の通ったストーリーを作り出し、それが文章理解を深めるといわれる。

文章を読んだあと、内容を知らない聞き手へ説明するためには、さまざまなスキルが必要になる。その点において再話は、相手に発話するという統語処理を通じてプロセスを考える力、批判的思考、コミュニケーション力、独学力を高める(Brown & Cambourne, 1990)。また、再話は英文の語彙習得にも効果があり(Morrow, 1985)、読書後の活動に取り入れる利点は大きい。

3.6 多読におけるプレリーディングとポストリーディング活動

多読学習でのプレリーディング活動の必要性はある程度認識されている。例えば、多読の代表的な教材であり英国の小学校で国語学習副読本と位置づけられている Oxford Reading Tree(初級レベル Stage1-4)の内表紙には、読書前活動として “Before reading: Look at the cover and read the title together.” という指示文が必ず記載され、表紙の絵とタイトルを見るプレリーディング活動が紹介されている。これも視覚を使って学習者のスキーマを活性化させる意図がある。だがこのシンプルな指示文だけで学習者の背景知識を活性化させることは難しい。また、多読は「学習者中心の読書活動」(酒井, 2005)と位置づけられ、プレリーディングをするかどうかは、読み手に任されているのが一般的である。多読の意欲を高め、深い理解を体感させるため、より具体的な手順を示したプレリーディング活動が必要である。

多読後のポストリーディング活動に要約や再話などの統語処理を取り入れる事例もある。読み手の理解度を深めて読書意欲を高めることが目的である。これらの統語処理活動の利点も、読解後の活動として広く認められている。ただし多読後の実施には賛成意見ばかりではない。授業内多読中に長文の要約を課す場合、学習者は書くことに囚われ、授業内での多読時間が減少することが指摘されている(高瀬, 2009)。

確かに、要約が負担となり、授業内多読の読書時間が減ることは、一見すると「読書を楽しむ」という多読の理念に反するようにも見える。し

かし、ポストリーディング活動に工夫を凝らし、達成感や読書意欲を向上させる負荷の軽い要約や再話活動を行ったとしたらどうであろうか。授業内多読の読書時間が減ったとしても、学習者は本に対する深い理解を体験し、違った視点で多読をより楽しめるであろう。そして、読書に対する目が開かれた学習者が、授業が終わった後に、自ら洋書を手に取ることも期待できる。多読におけるポストリーディング活動を否定的にとらえず、自立した読み手を作るための手段と位置づけることも可能であろう。

3.7 4技能統合型リーディングと多読

近年の学校英語教育では多技能統合型の指導が求められている。高等学校学習指導要領(外国语)「コミュニケーション英語Ⅰの思考力・表現力・判断力等の項目では、「日常的な話題や社会的な話題について、伝える内容を整理し、英語で話したり書いたりして、要点や意図などを明確にしながら、情報や自分自身の考えなどを伝え合うこと。」(p163)と記載されている。こうした視点に立つと、多読学習も「楽しみのための読書」という位置づけを大切にしつつも、その良質な英語のインプットを起点にして、スピーキングやライティングなど4技能統合型の学習活動に展開することは重要である。

4 多読学習支援ワークシート (Odyssey Reading Worksheet=ODR) 開発の背景

多読授業の課題や先行研究を踏まえ、多読をプレリーディング、インリーディング、ポストリーディングの3つに分割し、それぞれに明確な意図を持つ活動を取り入れることを検討した。

まず、プレリーディングでは、トップダウン処理を促して背景知識を活性化させる。これにより読書意欲を引き出し、本への関心を高める。生徒の多読意欲を引き出し、50分間集中して多読に取り組ませるための仕掛けである。

合わせて内容の推測を行い、本文の大意把握や理解を促す。前述のように、英文の理解には読書過程においてトップダウン処理とボトムアップ

処理の相互作用が必要である。だが英語への苦手意識を持つ学習者は、英文の細部にこだわるボトムアップ処理にとどまり、内容理解が妨げられ読書スピードも遅い。そこでトップダウン処理を促すような活動を取り入れ、大意を把握する力を伸ばして学習者の理解を深化させる。

インリーディングでは、読書を行う。漫然と多読するのではなく、意図や目的を持たせることで、読書への集中力を向上させる。

ポストリーディングではサマリーや再話などの統語処理活動を取り入れる。自分の頭の中で読書体験を再構築させ、文章理解を深めることを意図している。また読書後にアウトプット活動を設定することで、多読中に英文をより注意深く読むことを促す。

ODRでは上記の通り多読を3つのパートに分割し、更にリーディング以外にライティングやスピーキング、リスニングといった多技能活動を取り入れた。これまで多読において、読書体験の要約や再話などのライティングやスピーキングを取り入れた4技能統合型の実践報告例は少ない。それは授業内多読の目的の1つが、「読書時間の確保」であり、ライティングやスピーキングの統語処理活動を取り入れると、学習者の読書時間が減り、読書語数や読書冊数の減少をもたらす懸念があるためである。そこで学習者の読書量とアウトプット機会の確保を両立できるような多読授業の型を検討した。

以上の点を踏まえ、プレリーディングでの生徒の興味を喚起させる「トップダウン処理」と、ポストリーディングでのサマリーライティングやペア活動などの「統語処理活動」を取り入れた4技能統合型の多読支援用ワークシートOdyssey Reading Worksheet=ODRを開発した。これらの活動を網羅して標準化された多読用ワークシートがあれば、生徒にとって効果的な多読学習を行うための支援ツールになる。また教師側にとっても多読指導が容易になるとを考えた。以上がODR開発の背景である(写真3,4)。

Odyssey Reading Worksheet

No. _____ Date _____ / Year 20_____

1 Pre reading (3 min)

Title _____

Type of Writing :

Author/Series _____

Narrative (お話)

Word counts _____

Informative (情報や説明)

Time(min) _____

Argumentative(主張・根拠・結論)

1 What words will be in this book?

本の中で使われている、と予想したボキャブラリーを書き出しましょう（予想：黒→実際：赤）

2 What will the story be about?

タイトルと表紙から内容を予測しましょう（英語1～2文）

3 Make questions about this book.

質問を作成しましょう（英語1～2文）

この本を読む目的は何ですか？

<設定について><話の筋><登場人物><問題と解決策など>…

Sample questions What is the story about? / What is the most important message?

What does ●●● mean in the story? / What is the main character like?

5W1H (When Where Who What Why How) etc...

Class(-) No.() Name _____

■写真3: Odyssey Reading Worksheet 表

2 Post reading (5 min)**1 What connection do you have with this book?** 本とあなたの関連性

--	--

2 Who do you recommend this book to read ? How do you evaluate this book? 評価

☆☆☆☆☆	
-------	--

3 What is your favorite sentence in this book? お気に入りの一文

--	--

Writing Write the summary of this book. 本のサマリーを書きましょう

N <input type="checkbox"/> Setting 状況設定（場所、時代など） I/A <input type="checkbox"/> ()	N <input type="checkbox"/> Plot 話の筋（問題など） I/A <input type="checkbox"/> ()
Main idea (一言でいうとどんな話？)	
N <input type="checkbox"/> Character 登場人物 I/A <input type="checkbox"/> () N <input type="checkbox"/> Solution 解決 I/A <input type="checkbox"/> ()	

Speaking Visualize the story

相手のサマリーを聞いて、内容を絵で表現しましょう

--	--

Reflection Odyssey Reading

リーディングを行った後、ちょっとした気付きや変化を書きましょう

■写真4: Odyssey Reading Worksheet 裏

5

本研究における リサーチクエスチョン

本研究の仮説は、「ODRは多読学習の支援ツールとなりうる」である。この仮説を検証するため、設定したリサーチクエスチョンは以下の通りである。

RQ1

ODRの使用者(処置群)は、ODR未使用者(対照群)に比べて英語力を伸ばすか。

RQ2

ODRの使用者は、読書スピードを伸ばすか。

RQ3

4技能統合型のODRを取り入れた場合、学習者の多読語数・多読冊数はどう変化するか。

RQ4

進路多様高校のODR使用者は、入学後からの2年9か月間でどの程度多読語数・多読冊数を読めるか。また、どの程度GTECスコアを伸ばすか。

RQ5

ODRの使用者の多読への姿勢は、どのように変化するか。

RQ6

ODRの使用者は、ODRのプレリーディング活動とポストリーディング活動をポジティブに認識するか。

6

Odyssey Reading Worksheet(ODR)多読の実践

6.1

Odyssey Reading Worksheet (ODR)実践の概要

6.1.1. 予備調査とワークシートのプロトタイプの開発

本実践を行うにあたり、ワークシートのプロトタイプ(原型)の作成と検証を目的に予備調査を実施した。予備調査では、多読ワークシートをコミュニケーション英語Ⅱの50分授業5回で使

用した(2019年2月)。試用期間中、生徒からアンケートを実施し、ODRの表現や構成を調整した。改良を重ね、2019年5月にワークシートを完成させた。

6.1.2. Odyssey Reading Worksheet (ODR)

『Odyssey』は古代ギリシャの長編叙事詩である。ギリシャ軍のトロイ攻略後の帰国物語で、知将オデュッセウスの苦難に満ちた「放浪冒険旅行」だ。この『Odyssey』に倣い、多読を少しでも知的でエキサイティングな冒険として楽しんでほしい。そこで多読用ワークシートを“Odyssey Reading Worksheet”(ODR)と命名した。

ODRは、プレリーディング活動で内容想起(3分)→インリーディングで読書→ポストリーディングで内容をまとめてアウトプット(5分~8分)、という一連の流れで多読を行う。ワークシートの表面はプレリーディング用、裏面はポストリーディング用とした。

6.1.3. 対象授業と対象者

本実践はコミュニケーション英語Ⅲ(週5時間)中、週2時間(多読1時間・多聴1時間)の授業において、研究者が担当するクラスで行った。習熟度別授業の発展クラス20名×3クラスが対象である。授業の実践場所はいずれも特別教室であり、多読授業は図書室、多聴授業はCALL教室で行った。

実施期間は2019年9月1日~11月30日の約3か月である。その間、ODRをいずれのクラスでも、合計27回実施(内訳:多読授業14回、多聴授業13回)した。本実践の対象者は3年生117名である。そのうち研究者が担当するクラスの生徒(n=60)を処置群、その他のクラスの生徒(n=57)を未使用者(対照群)とした。

6.1.4. Odyssey Reading の手順(資料2)

(1)選書(3分)…推奨多読リスト(資料1)を参考に、生徒自身が書架から好きな本を選ぶ。3分間で3冊程度を目安にする。(写真5,6)

(2)プレリーディング(3~5分)…多読前の準備活動である。本の種類分析や既存の背景知識を活性化、内容を推測する活動を行う(写真7)。

(活動:Reading／Writing)

- ① type of structure選択:本の種類を Narrative(物語型), Informative(情報型), Argumentative(論文型)から選び、ボックスに✓を入れる。
- ② 予測型のセマンティック・マッピング(写真10):どのような語彙がテキスト内で使われているかを予測する。タイトルと挿絵をよく見て、本の中で使用されていると予想した単語を、1分間で出来るだけ多く書き出す。想起できた単語数を記録する。マインドマップ状に、スピードを重視して書かせる(まだ本の中身は見ない)。
- ③ 内容推測:タイトルと挿絵から、本の内容を推測し仮説を立てる。推測した内容は英語1~2文で書く。推測する活動自体を大切にし、正確に拘りすぎないよう伝える(写真11)。
- ④ 質問作成:本の内容についての質問を作成し、読書の目的を作る。質問を作ることが難しい場合、5W1Hなどサンプルの質問を参考にする(写真12)。
- (3) インリーディング(10~40分)…ここで初めて洋書を読み始める。多読3原則に沿って多読を楽しむ。生徒の状態や授業時間に合わせてインリーディング時間を調整する。この部分は普通の多読と同じである(活動:Reading)。
- (4) ポストリーディング(5~8分)…読書後の振り返りである。プレリーディング活動を振り返り、読んだ本の内容のサマリー・ライティングや、要点をペアと伝えあうスピーキングなどの統語処理活動を行う(活動:Writing／Speaking／Listening)。
 - ① セマンティック・マッピングの確認:読書後、プレリーディングで想起した単語と実際の単語を比較する。想起していた単語が本文で使われていた場合、赤ペンで○を囲む。また重要と思われる単語を赤ペンで追記する。マッピングに単語を追加し、本の内容を構造化する。
 - ② 作成した質問への回答:読了後に質問への答えを書き出す。
 - ③ 関連づけ:読書後、自分と本の内容の「共通点」を見つける。洋書との自己関連性を高め、記

憶に残りやすくする(写真13)。

- ④ 評価:誰におすすめしたいかという視点で本を評価する。内容を客観的に認識することを促す。
- ⑤ お気に入りの1文:コピーイングである。お気に入りの英文を書き写す。ライティングや会話で使ってみたいという視点から探す。
- ⑥ サマリー・ライティング:サマリー作成のための代表的な4つの視点から、要約を作成する(写真14)。まず、ボックスの真ん中に、要約の大項目にあたる Main idea を書き、何の話であるかを一言で表す。次に type of structure を選び、N(Narrative) or I/A(Informative/Argumentative) のどちらかに✓する。Narrative の場合、setting, plot, characters, solution(要約の中項目)を✓する。Informative/Argumentativeの場合または setting, plot, characters, solution の視点を用いるとき約しにくい場合、I/Aに✓し、()に要約の中項目を自ら考えて記入する。ボックスごとの要約の中項目の視点から、本文の要約を書く。サマリーは単語やキーワードではなく、主語+述部(動詞と目的語や補語、修飾語)で書く。
- ※ここまでで1冊読み終えた学習者は、別の本で多読を続ける。
- ⑦ スピーキング活動(授業終了5分ほど前に開始):インフォメーションギャップを利用したスピーキングとリスニング活動である。まず、ペアを作り、相手とODRを交換する。生徒Aは、サマリーをもとに、1分間英語で本の要約を話す。生徒Bは相手の英語を聞き取りながら、相手のODRへ絵を描き、相手の発話をヴィジュアル化する。AとBの役割を交換する(写真8, 9)。
- ⑧ 振り返り:ODRを使った多読活動を振り返る。気づきや自分の変化を書き、ODR多読を終える。

以上が Odyssey Reading Worksheet 多読の流れである。

Odyssey Reading Worksheet多読の流れ

■写真5



(1)選書(3分)

■写真6



■写真7



(2)プレリーディング 3~5分 ⇒
(3)インリーディング

■写真8



(4)ポストリーディング(スピーキングとライティング)5~8分

■写真9



6.1.5.定型化されたOdyssey Reading Worksheet多読授業の利点

上述のように、Odyssey Readingは4技能統合型の多読授業手順をワークシート上に標準化したものである。その利点の1つは、標準的な多読授業の再現性が向上する点である。通常の多読授業と比べて授業での活動の安定性が高い。ワークシートを基盤にして、リーディング以外のスピーキングやライティング、リスニング活動を包含した多読授業を容易に行うことができる。教師側にとって、多読指導経験の有無にかかわらず、ODRを使用することで多読指導が容易になるであろう。

2つ目の利点として、多読授業の手順を可視化したことにより、多様な背景の学習者にも活用できることが挙げられる。多読手順を可視化したワー-

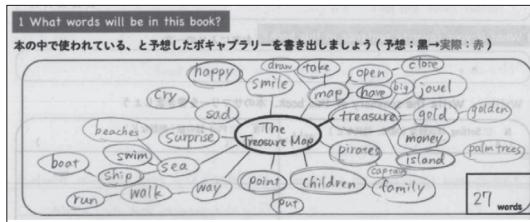
クシートを使用すると、プレリーディングやポストリーディング時の指示が具体的な活動に落とし込まれているので、学習者の本への関心を高め、多読への集中力を維持することが期待できる。

6.2 Odyssey Reading Worksheet 多読の調査手順

6.2.1.データ収集

本研究の実施にあたり、生徒の英語力伸長を定量的に測定するため、ODR使用者(処置群)と未使用者(対照群)は GTEC for STUDENTS を受検する。実施時期は ODR導入前の2019年7月と ODR実施後の12月であり、前者をプレテスト、後者をポストテストと設定する。リーディングスコアとリーディング速度の WPM (Word Per

■写真10:セマンティック・マッピング



■写真13:関連づけ,評価,お気に入りの一文

2 Post reading (5 min)	
1 What connection do you have with this book? 本とあなたの関連性 I want to find the treasure box in the future.	
2 Who do you recommend this book to read? How do you evaluate this book? 評価 I wonder about last page but pirates are cute. I like this book. ★★★★☆	
3 What is your favorite sentence in this book? お気に入りの一文 "His ship is here. Now he can get his treasure back."	

■写真11:内容推測

2 What will the story be about? タイトルと表紙から内容を予測しましょう（英語1～2文） This story is about treasure. A man have a treasure map, he and his friends find the treasure box.	
---	--

■写真12:質問作成

3 Make questions about this book. 質問を作成しましょう（英語1～2文） Questions What is the story about? Answers It's about two kids go to a party and dance.	
---	--

■写真14:サマリーライティング,スピーキング

Writing Write the summary of this book. 本のサマリーを書きましょう	
N <input type="checkbox"/> Setting 状況設定（場所、時代など） 1/A□()	N <input type="checkbox"/> Plot 話の筋（問題など） 1/A□()
Pirates and Children are at island.	The Magic Key take children beach.
	They met pirates, and they helped pirates for find treasure.
Main idea（一貫でいとどな話？） King's Treasure Children, Man, Pirates and Big Jim, King Anneena caught place there is a treasure box in the beach, so pirates can find it but King's ship will arrive at the beach, the treasure back for him.	
N <input type="checkbox"/> Character 登場人物 1/A□()	N <input type="checkbox"/> Solution 解決 1/A□()
Speaking Visualize the story 相手のサマリーを聞いて、内容を絵で表現しましょう	
Reflection Odyssey Reading リーディングを行った後、ちょっとした気付きや変化を書きましょう	

■写真15:Odyssey Reading Worksheet (ODR)の生徒記入例

Odyssey Reading Worksheet	
1 Pre reading (5 min)	
Title New Year's Day Date 1/23 Year 2020	
Author/Series David F. Moore Type of Writing : □Narrative (お話) □Informative (情報や説明) □Argumentative (主張・根拠・結論)	
Word counts 283 Time(min) 5.6	
1 What words will be in this book? 本の中で使われている、と予想したボキャブラリーを書き出しましょう（予想：黒→実際：赤）	
2 What will the story be about? タイトルと表紙から内容を予測しましょう（英語1～2文） I think this book about many countries' January 1 life.	
3 Make questions about this book. 質問を作成しましょう（英語1～2文） Questions How many about country? Why had a this book? Answers Maybe 8 country. Because food is cute and good taste.	
Sample questions What is the story about? / What is the most important message? What does ●●● mean in the story? / What is the main character like? SWIH (When Where Who What Why How) etc...	
2 Post reading (5 min)	
1 What connection do you have with this book? 本とあなたの関連性 I'm in house with Shirley Johnson 1.	
2 Who do you recommend this book to read? How do you evaluate this book? 評価 So interesting. Because team help country culture the book. ★★★★☆	
3 What is your favorite sentence in this book? お気に入りの一文 "People celebrate many new things"	
Writing Write the summary of this book. 本のサマリーを書きましょう	
N <input type="checkbox"/> Setting 状況設定（場所、時代など） 1/A□()	N <input type="checkbox"/> Plot 話の筋（問題など） 1/A□()
This book about many country's people life in January 1.	I think book mostly about its book. Because twelve different people, they are doctor and go to city. So, book mostly draw Japanese title and other country's culture.
	I hope book mostly talk in Japanese.
Main idea（一貫でいとどな話？） Happy New Year!	
N <input type="checkbox"/> Character 登場人物 1/A□()	N <input type="checkbox"/> Solution 解決 1/A□()
Speaking Visualize the story 相手のサマリーを聞いて、内容を絵で表現しましょう	
Reflection Odyssey Reading リーディングを行った後、ちょっとした気付きや変化を書きましょう	

Minutes) の変化量を検証する。

また、ODR 実施により多読読書語数と冊数の変化を見るため、ODR 多読実施期間(2019年9月～11月)と、実施前(2019年4月～6月)の前後3か月ずつでの多読読書語数を比較する。

ODR を使用した学習者が、高校3年間でどれだけの洋書を多読し、英語力を伸長させたかも測定する。そのために高校入学時から卒業前までの累計多読語数と冊数、同時に、GTEC スコアの高校1年入学時から高校3年12月までの変化量を測定する。

生徒の多読への情意面の変化を把握するため、アンケートを実施する(資料3)。実施時期は2019年9月と12月の2回であり、前者をプレテスト、後者をポストテストとした。加えてODRへの生徒の認識を理解するため、自由記述式のアンケートも2019年12月に行った(資料4)。

6.2.2. データ処理

RQ1では、GTEC for STUDENTS のスコアを使用する。吉田(2006)の推奨する分析方法として、処置群と対照群のポストテストからプレテストの点数を引いた変化量の平均値に対してt検定を行う。

RQ2では、読書スピードの変化を検証するため、WPM(Word Per Minutes)を測定する。WPMは読解読書語数であり、word ÷ 読解に要した時間(分)に、内容理解度(内容に関する設問の正解率)を乗じた値と定義する。GTEC for STUDENTS の結果を用いる。ODR の実施前と実施後での変化量を算出し、t検定を実施する。

RQ3では、ODR 使用期間中の多読読書語数と冊数を、ODR使用者(処置群)と未使用者(対照群)の平均値で比較する。

RQ4では、多読多聴手帳から得られた累計多読語数と冊数について、基本統計量(平均・最大・最小・標準偏差)を求める。多読語数と英語力の関係を捉えるため、GTEC スコアと多読語数を散布図で分析する。

RQ5では、ODR 多読活動を通じ、学習者の多読意識がどのように変化したかを調べるために、アンケート項目(資料3)に対して、ODR 実践前と実践後の平均点の差が統計的に有意か確かめるため、有意水準5%で両側検定のt検定を実施する。

RQ6では、ODR 多読の使用者60名に対して、ODR 多読実践後(2019年12月)にODRへの評価を調査する。調査にはアンケート用紙を使用した。質問内容は35項目について5件法で回答させる。なお、アンケートに加えて自由記述式も実施する。

7 研究結果

7.1 RQ1: ODR の使用者(処置群)は、ODR 未使用者(対照群)に比べて英語力を伸ばすか。

7.1.1. 結果

ODR使用者(処置群 = β クラス)と未使用者(対照群 = α クラス)の2群に分け、GTEC for STUDENTS のReading スコアの変化量(ODR 多読前とODR 多読後)の平均値を算出した。処置群のリーディングスコアは平均で177.08から184.20へ上昇した。対照群の平均値は153.07から148.98へ下がった。平均値の変化量は処置群で7.12、対照群で-4.09となった(表2)。この2つの平均値の差が統計的に意味のある差かどうか検定するためt検定を行ったところ、統計的に有意な差が見いだされ($t=2.75$, $df=115$, $p<0.01$)、ODR 多読を実施するかどうかがReading スコアの伸長に関係していることが分かった。

■表2: GTEC Readingスコアの平均値とODR実施前後の変化量

	ODR前 (2019年7月)	ODR後 (2019年12月)	変化量
$\beta=60$	177.08	184.20	7.12
$\alpha=57$	153.07	148.98	-4.09

7.1.2. 考察

プレリーディングでのトップダウン処理を促進することで、学習者のスキーマが活性化され、効果的な読みに至る可能性が示された。また、ポストリーディングの再話やサマリー作成などの統語処理を行うと、頭の中で英文内容を再構築し、文章理解を深めることができると示唆される結果となった。

一方でODRを使用しなかった対照群のリ-

ディングスコアが -4.09 低下した。これは2年半に及ぶ多読学習への「飽き」や、高校3年次の早期進路決定に伴う学習意欲の減退といった背景が考えられる。同様の条件はリーディングスコアの上昇を示した処置群にも当てはまるため、4技能統合型多読であるODRの使用は学習者の多読意欲を刺激したことが考えられる。

7.2 RQ2:ODRの使用者は、 読書スピードを伸ばすか。

7.2.1. 結果

ODR使用者(処置群 = β クラス)に対し、ODR 多読前と ODR 多読後の平均 WPM を集計したと

ころ、ODR 多読前で 80.5、ODR 多読後で 85.2 となかった。ODR 未使用者(対照群 = α クラス)の平均 WPM は 69.2 から 65.8 に下がった(表3)。

ODR 処置群と ODR 対照群の2群に分け、WPM 値の変化量(ODR 多読前と ODR 多読後)の平均値を算出したところ、ODR 処置群で 4.7、対照群で -3.4 となった。この2つの平均値の差が統計的に意味のある差かどうか検証するため t 検定を行ったところ、統計的に有意な差が見いだされ($t=3.16$, $df=112$, $p<0.01$)、ODR 多読を実施するかどうかが WPM の変化に関係していることが分かった。

■表3: ODR実施前と実施後の平均WPM値

クラス	WPM① 2019/7 (a)	最大値 (b)	標準偏差 (c)	WPM② 2019/12 (d)	最大値 (e)	標準偏差 (f)	変化量 (d-a)	人数
β	80.5	140.0	22.0	85.2	140.0	21.3	4.7	60
α	69.2	95.0	10.3	65.8	95.0	10.3	-3.4	57

7.2.2. 考察

ODR 処置群の平均 WPM が対照群に比べて有意な上昇が認められた。トップダウン処理が英文理解に有効とされた Bransford&Johnson (1972) の結果にも一致する結果となった。このことから ODR は学習者の背景知識を活性化させるトップダウン処理を促す効果があることが示唆された。特に、ポストリーディングでのスピーキング活動が設定されており、学習者は本の内容を理解して相手に伝える必要があるため、学習者が英文を正確かつ素早く読む動機付けにもなっている。結果として WPM の上昇を示したといえる。一方で ODR 対照群の WPM が低下した背景は、多読学習への飽きといった意欲面の低下が考えられる。

7.3 RQ3:4技能統合型のODRを 取り入れた場合、学習者の多読語数・ 多読冊数はどう変化するか。

7.3.1. 結果

ODR を実施した処置群(β)の3か月間(2019年9月～11月)の多読総語数は、平均で 36,314 語、最

大 724,862 語、標準偏差 91,223.7 であった。また多読冊数は平均で 72 冊、最大 321 冊、標準偏差 55.9 であった。一方、ODR 実施前の3か月(2019年4月～6月)の多読語数は平均で 18,129 語、最大 85,608 語、標準偏差 19,158.5 であった。多読冊数は平均で 44 冊、最大 400 冊、標準偏差 59.5 であった。ODR 実施期間と未実施期間を比較すると、平均語数は 18,185 語増加し、平均冊数は 28 冊増加した。(表4)。

ODR を実施しなかった対照群(α)では、3か月間(2019年9月～11月)の多読語数は平均で 9,505 語、最大 66,319 語、標準偏差 13,661.4 であった。多読冊数は平均 24 冊、最大 293 冊、標準偏差 46.4 であった。一方、ODR 実施前の3か月(2019年4月～6月)の多読総語数は、平均で 14,145 語、最大 122,931 語、標準偏差 21,138.5 であった。多読冊数は平均 39 冊、最大 514 冊、標準偏差 70.3 であった。対象期間での多読平均語数と冊数の差異は、ODR 実施前と実施後で、平均で 4,640 語、15 冊の減少を示した。

■表4: ODR実施期間と未実施期間の多読読書語数・冊数の比較

①未実施期間(2019.4-6)と②ODR実施期間(2019.9-11)の多読語数・冊数比較

語数	①2019年4月-6月			②2019年9月-11月			平均の差 ②-①
	平均	最大	標準偏差	平均	最大	標準偏差	
β	18,129	85,608	19,158.5	36,314	724,862	91,223.7	18,185
α	14,145	122,931	21,138.5	9,505	66,319	13,661.4	-4,640

冊数	①2019年4月-6月			②2019年9月-11月			平均の差 ②-①
	平均	最大	標準偏差	平均	最大	標準偏差	
β	44	400	59.5	72	321	55.9	28
α	39	514	70.3	24	293	46.4	-15

7.3.2. 考察

高校3年生の9月～11月は特に多読学習意欲の減退する時期である。ODR多読の対照群(α)では、想定通り多読読書語数と冊数が減少した。これは多読学習への飽きだけでなく、勤務校の生徒80%以上が、AO入試を利用して早期に進路を決定し学習への動機が薄れたためである。

一方でODR処置群(β)において、授業内の多読時間が減ったにも関わらず、多読読書語数と冊数の増加が見られた。ODR多読は4技能統合型の多読授業のため、スピーフィングやライティングにも時間を割くので、学習者が洋書を多読する時間自体は減る。そのため多読量も減少することが想定された。しかし、結果はODR多読時期が未実施時期の平均累計読書語数と冊数を上回った。

このことからトップダウン処理が読み手の関心を高める利点があると報告された卯城(1993)と一致する結果となった。以上のことから、ODRによって「読みの変化」や「4技能活動」を感じた学習者が、多読や英語学習への意欲を高め、授業外でも多読用図書を借りるなど、行動変容をもたらす可能性が高いことが示唆される結果となった。

7.4

RQ4: 進路多様高校のODR使用者は、入学後からの2年9か月間でどの程度多読語数・多読冊数を読めるか。またどの程度GTECスコアを伸ばすか。

7.4.1. 結果

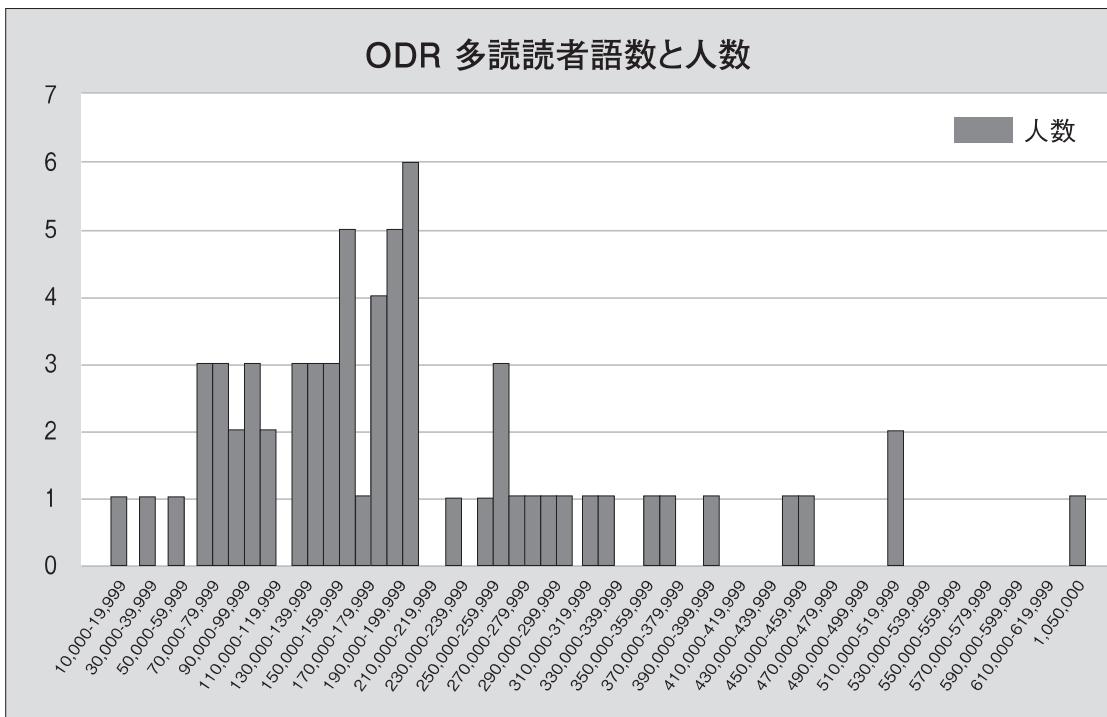
ODR多読の処置群60名の2年9か月間の最終的な多読総語数は、平均で203,355語であり、最高は1,095,685語であった。また、読了冊数については平均で639冊となり、最高で1,741冊であった。表5と図1は上記の結果である。

ODR処置群60名のGTEC平均スコアは、初回

■表5: ODR多読処置群の多読読書語数と冊数

高校3年12月(2019年)

	2019年12月	
	多読語数	多読冊数
人数	60	
平均	203,355	639
最大	1,095,685	1,741
最小	8,144	112
中央値	178,401	619



■図1: ODR処置群の3年間の多読読書語数ヒストグラム

受検の高校1年次4月から高校3年次12月の間で, Reading で131.6から184.2に上昇した。Listening は153.5から195.8へ, Writing は187.1から218.0に達した。WPM (Word Per Minutes) は56.8から85.2へ伸びた。CEFR-J レベルは Reading と Listening で A1.3 から A2.2 へ, Writing は A1.3 から A2.1 へ伸長した(表6)。

全国平均値を見ると, 高校1年の全国平均は Reading 152, Listening 159, Writing 196であり, 対象生徒のスコアはいずれも全国平均を下回っていた。しかし, 高校3年次の全国平均値は Reading 176, Listening 183, Writing 203であり, 多読多聴学習を行った学習者はすべてにおいて全国平均を上回った。累計多読語数と

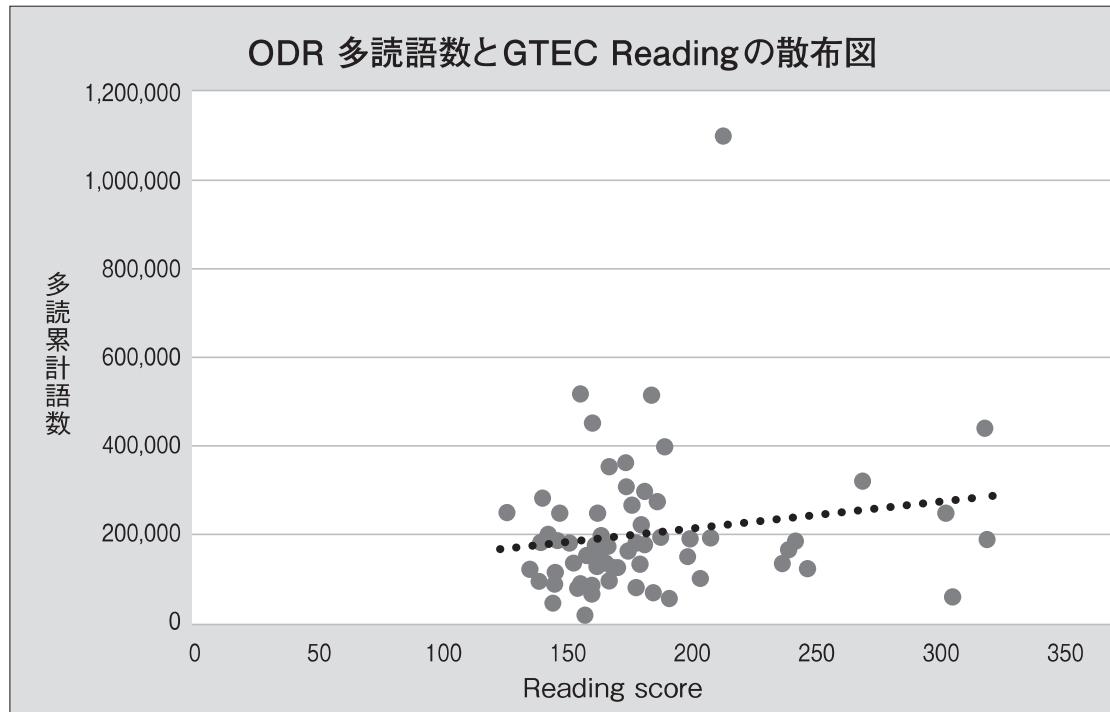
GTEC スコアの散布図は以下の通りである(図2)。

7.4.2. 考察

上記の結果から, やさしい英語の絵本を約600冊, 20万語程度を多読多聴した学習者は, GTEC スコアは A1.3 から A2.2 付近まで伸長していることが示された。全国平均値を比較すると, 高校1年次は3技能いずれも全国平均以下であったが, 高校3年次にはすべてにおいて全国平均を上回る数値に達した。これらは多読原則に沿った多読の効果を提唱している酒井(2002), 古川(2010)らの結果と一致するものとなった。このことから, 中学英語の基礎が未定着の高校生に対して多読多聴を取り入れた英語教育の有効性が示唆され

■表6: ODR処置群のGTECスコアの推移

GTEC	2017年4月 (高校1年次)	CEFR-J	2019年12月 (高校3年次)	CEFR-J	変化量
Total	472.2		598.1		125.9
Reading	131.6	A1.3	184.2	A2.2	52.6
Listening	153.5	A1.3	195.8	A2.2	42.3
Writing	187.1	A1.3	218.0	A2.1	31.0
WPM	56.8		85.2		28.4



■図2: ODR処置群の多読読書語数とGTEC Readingスコアの散布図

る結果となった。

ODRの実施期間は2年9か月中、最後の3か月間と限定された期間であったが、学習意欲の低下する3年の後期に、リーディングスコアや速読値を伸長させるなど英語力を引き続き伸ばした。ODR実施期間中に読書量を増やし、多読読書語数を100万語まで伸ばした生徒も出現した。これはODRという多読ツールの導入により、多読学習に変化と刺激が生じ、対象者は単調になりがちな多読学習に対して集中力を維持できた結果といえよう。

7.5 RQ5:ODRの使用者の多読への姿勢は、どのように変化するか。

7.5.1. 結果

RQ1で英語力の変化も確認されたが、テストで測定できる以外の部分にどのような変化が生じているかを分析することも重要である。ここでODRを使用した学習者の情意面の変化を概観する。p値については*:5%有意を表す。

調査した11項目中、統計的に有意な変化が確認されたのは3項目であった。「3. 多読用の本を読む前に、絵やタイトルから内容を推測してい

る」(p= 0.0014), 「4. 多読の時、「速読」を意識して読んでいる」(p=0.0144) は有意な上昇を示した。また、「8. 多読の時、意味の分からない単語は、ひとつ一つの意味を考えたりせず、飛ばして読んでいる」(P=0.0379) は有意に減少した(表7)。

統計的な有意差は確認されなかったが、注目すべき点もある。まず意欲面であるが、「1. 多読は楽しい」, 「2. 多聴は楽しい」において、ODR実施後に数値の上昇傾向がみられた。他にも多読本の読み方の変化が推察できるのは「5. 多読の時、知っている単語の意味をつなぎ合わせて、なんとなく読んでいる」, 「10. 多読の時、英文の主語や動詞など、文の構造を意識して読んでいる」, 「11. 多読の時、英文を日本語に訳して読んでいる」の3つの項目である。いずれも統計的に有意な変化ではないが、ODR多読実践後、知っている単語をつなげてなんとなく読んだり、英文の構造を意識し過ぎたり、英文を日本語へ翻訳する読み方が減少傾向を示した(表7)。

7.5.2. 考察

検定結果から読み取れることは以下の通りである。まずODRを使用した学習者は、使用前と

比較して、読書前の推測活動を行うようになつた。また速読の意識も高まつた。トップダウン処理により背景知識を使いながら読書をするので理解が容易になり、結果として読書スピードが上がつたことが示唆された。また未知の単語を飛ばし読みする学習者が少なくなったことで、学習者はより正確に英文を捉えるようになった。

このことからポストリーディングの統語処理は、学習者が英文を深く捉える手助けをする可能性があることが判明した。

統計的な有意差は確認できなつたが、アンケートからは多読への意欲を高める傾向が示され、中学英語の基礎が未定着の生徒にとって、これらは多読を楽しむ上で好ましいといえよう。

■表7: 多読アンケートデータ

Q	事前アンケート		事後アンケート		変化量	t値	p値	*p<.05
	平均	標準偏差	平均	標準偏差				
1	3.344	0.930	3.391	0.866	0.047	-0.319	0.7505	
2	3.548	1.035	3.581	0.897	0.032	-0.207	0.8368	
3	3.274	0.961	3.790	0.890	0.516	-3.355	0.0014	*
4	2.742	0.974	3.145	1.053	0.403	-2.519	0.0144	*
5	3.839	0.995	3.806	1.006	-0.032	0.209	0.8350	
6	3.226	1.165	3.306	1.154	0.081	-0.394	0.6951	
7	3.887	0.870	3.710	0.998	-0.177	1.055	0.2958	
8	3.806	1.114	3.355	1.189	-0.452	2.122	0.0379	*
9	3.839	1.011	3.823	1.079	-0.016	0.085	0.9322	
10	2.532	1.097	2.484	0.987	-0.048	0.262	0.7941	
11	3.000	1.215	2.903	1.277	-0.097	0.574	0.5679	

7.6 RQ6:ODRの使用者は、ODRのプレリーディング活動とポストリーディング活動をポジティブに認識するか。

7.6.1. 結果と考察

35項目のアンケートについて、数値の上位項目から並べたものを記したところ(資料4)、質問35項目中、31項目で5件法中3.0以上となり、学習者はODRのプレ・ポストリーディング活動をポジティブに認識していることが示唆された。

第1に、内容理解の深化に関する項目で高い値を示した。まず、「1. 普通の多読に比べて、ODRを使って多読すると本の内容理解が深まる」(4.18)では、ODRは読書理解を深めるという認識をもつてゐることが示された。特に、「26. ODRのポストリーディング活動で、本の内容理解が深まる」(4.19)や、「17. プレリーディング活動(ODRの表面)をすると、読む本の理解が深まる」(4.02)、「27. ODRでペアとスピーキング活動をすると、本の

理解が増す(4.05)など、読書前のトップダウンアプローチと、読書後の統語処理活動は読書理解を深めると認識された。

第2に、多読の読み方に変容が見られた。「22. セマンティック・マッピングの時、タイトルや表紙の絵をよく観察するようになった」(4.11)、「13. ODRを使わないので多読するとき、タイトルや表紙をよく観察するようになった」(3.55)の示す通り、読書前に本のタイトルや表紙の絵を観察する読み方への変容が示された。また「15. プレリーディング活動(ODRの表面)をすると、多読の本が読みやすく感じる」(3.50)で、多読用図書を容易に読めるようになったと認識していた。

第3に、英語技能の向上への有用性が認識されていた。「7. ODRで多読多聴を行うと、普通の多読に比べて Writing が伸びると思う」(4.15)、「5. Reading が伸びると思う」(3.92)、「8. Speaking が伸びると思う」(3.84)、「6. Listening が伸びると思う」(3.29)など、ODR多読によって Writing>

Reading>Speaking>Listeningの順に英語力向上に有効だと認識された。

第4は、多読学習への情意面のポジティブな認識が見られた。「4. ODRを使うと、毎回の多読授業で達成感を得られる」(3.60),「2. 普通の多読よりも、ODRを使った多読は楽しい」(3.42)など、毎回の多読授業内で、達成感や楽しさを感じることが分かった。

逆に、低値を示した項目は次の通りである。「34. 高校卒業後も多読多聴を続けるつもりだ」(2.92),「35. もしODRをつかうことができれば、今後も一人で多読学習を続けることができると思う」(2.79),「31. Post readingの『自分とのかわり』を考えると、本文の理解が深まる」(2.61)などが示す通り、多読を楽しむ自律した学習者を育てるることは、今後の課題といえる。

7.6.2. 自由記述

ODRに関する自由記述アンケートを行った結果、いくつかの興味深いコメントがあった。第1に、多くの学習者はODRで多読時に「深い読み」を体感している。「読む前にどんな内容なのか自分で予測するから、その予測が間違っていた時、こういう内容だったのだ、と普段ODRなしで読んでいるときよりも感じられる。」「内容を推測することで、本を読み終わった後の内容と比べることができ、どこが違ったかが明確にわかるようになる。」「自分自身で内容を予測し、より興味を持った状態で多読に臨める。また、実際に読んだ後と前とで比べてみると内容がはっきりする。」プレリーディングで内容を推測することで、多読活動に驚きや発見が生まれ、深い内容理解を味わえた学習者が多く現れたといえる。

第2に、ODRによって学習者自身が創造性を發揮し、多読の楽しさや意欲を感じる学習者が現れた。「表紙を見て内容を想像してから本文を読むので、内容が理解しやすい。多読が苦手であったり好きではない人でも、集中して取り組めるし達成感があるので楽しめると思う。」「想像力が豊かになる。本を読む前の段階での活動が、非常に意味のある活動だと思う。」「自分の想像した内容と、読んでからの内容を比べることで内容理解が深まる。自由に想像できるから楽しい。」「推測力が身につく。自分が推測した通りの内容か、そうで

ないのか、多読していくでわくわくする。」「いろいろな感覚を使って多読を行うので、多読活動に積極的になる。」この自由記述から楽しみのための読書活動である多読学習と、学習者の創造性を刺激し知的な喜びを体感させるODRの両者は、親和性が高いといえよう。また「通常の多読に物足りなさを感じていたが、ODRが導入された高校3年は多読が楽しかった。」「高校1年次から取り組みたかった。」というコメントもあり、生徒自身もODRによる変化を体感していたことがうかがえる。

第3に、英語に苦手意識を持ち、多読に苦痛を感じていた学習者にとって、ODRは英文への抵抗感をやわらげ、本を読みやすくするツールともなっていた。「本を読む前に内容を想像してから読むことで、想像力がつくので、本を読むことへの苦手意識が薄れる。」「本を読むことが楽になる。」「多少わからなくとも、全体を通してなんとなく理解できるようになるので、英語に対する『難しい』という意識が薄れる。」「本の内容を予測してから多読すると、本を読んだときに内容が頭に入りやすい。」このような記述が多かったのはプレリーディングでの内容推測活動により背景知識が活性化され、英文を読む負担を軽減する効果があつたためであろう。

第4に、普通の多読と比較して、ODRの4技能統合型多読に価値を見出す学習者もいた。次のような記述をしている。「読むだけの活動だった多読が、ODRにより読む・話す・聞く・書くという4つの活動に展開できる点がよい。」「ODRは多読授業で4技能を使うので効率よく英語力が高まると思う。」「高校入学から2年間、ODRなしでの多読は少し物足りなさがあったが、3年になって導入されたODRのおかげで、4技能すべてにおいて向上を感じている。」「ペアで説明しあうことにより、リーディングだけでなくリスニングやスピーキングの練習になる。」「ODRの最後にペアでスピーキング活動することで、英語の表現力や英語の発想力が向上していることを実感できる点がよい。また英語で話すことへの恥やためらいがなくなり、自信をもって話すことが出来る点もよい。」これらのコメントから、ODRのを目指した4技能統合型多読の「型」を作るという試みは、ある程度成功したといえよう。

一方でODRの改善点に対するコメントは次の通りである。「ODRに時間がかかり、1回の授業で洋書を1~3冊程度しか読めない。」など、授業内の読書時間の減少についてのコメントが多くかった。また、授業時間内で終えるというODRの性質上、短めの本を選びがちになるという選書への影響もあった。「ODRはとても良い活動だが、長い本を読んでODRをやりたくても、授業時間内に終えられないのがつらい。その結果、短めの本を選んでしまう。」こうした時間的な制約に対し、「ODRの回数を少し減らしてもいいと思う。」といった実施回数の彈力的な運用に関する意見もあった。

最後に、絵本や多読用図書だけでなく、他の英文素材への応用を提案するコメントもあった。「洋書だけでなく、新聞や論文を読むときに使ってみたら面白いと思う。」このような更に難易度の高い英文を読むときに、ODRを使用したいという学習者の存在は、このワークシートの秘める可能性を示すものと言える。またライティングへの意欲の高まりを示す意見もあった。「本の要約だけではなく、本を読んでどう思ったか、自分の意見を書けるところがあるといい。」「ODRの要約もとてもよい活動だが、もっと長めの要約や自分の意見も書きたい。」等である。ODRが学習者のアウトプット意欲を刺激したことが垣間見える。

その一方で、トップダウン処理についての理解が不十分なままODRに取り組んだ生徒は、「タイトルと表紙から内容を予測することで何につながるのか、理論的な説明を知りたかった。」と述べた。ODRの導入時に理論面の説明を丁寧に行うことが必要であろう。

8

まとめと今後の研究課題

8.1 まとめ

本実践研究で明らかになったことは次の通りである。

(1) ODRで多読学習を行った学習者は、リーディング力を伸ばした。ODRを使用しなかった学習者は、逆にリーディングスコアを下げたため、

ODRはリーディング力の向上とともに、多読学習で起こりうる「飽き」や「マンネリ」への刺激になる可能性がある。

- (2) ODR使用者は未使用者よりも読書スピードの伸長が見られた。ODRが学習者の背景知識を活性化させるトップダウン処理を促す効果を示した。
- (3) ODR実施期間と実施前の期間を比較したところ、ODR実施期間の多読読書語数・冊数は増加した。これはODRによって「読みの変化」や、4技能統合多読を体感した学習者が、多読学習への意欲を高め、授業外でも多読用図書を借りるなど、ODRの使用が学習者の行動変容につながった可能性がある。
- (4) 高校生活3年間でやさしい英語絵本を平均で約600冊、20万語程度多読していた生徒は、GTECの3技能スコアはA1.3からA2.2程度まで伸長が見られた。多読多聴学習は中学英語の基礎が未定着の高校生に対して一定の効果が見られた。
- (5) プレ・ポストアンケートから、ODR使用者は読書前の推測活動を行うようになり、かつ速読の意識を高めた。トップダウン処理による背景知識を使用して読書するため、理解が容易になり英文の読書負担が軽減され、読書スピードも向上したといえる。また、ODR使用により、意味の分からない単語に出会ったとき、意味を考えず読み飛ばすことが減少した。これは再話やサマリー作成などのポストリーディング活動の統語処理を通じ、学習者が英文を深く捉える可能性を示唆した。加えて統計的な有意差は確認できなかったが、ODR使用によって多読は楽しいという認識が増えた。
- (6) ODRに関するアンケート結果から、学習者はODRによって英文の内容理解が深まるなどポジティブな情意面の変化を体感していた。また、自分の「読み方」の変化や、想像力を刺激される点、リーディング以外の多技能の向上を実感していることがわかった。
- (7) 自由記述のアンケートで、ODRによって授業内の読書時間が減ったと感じる学習者が一定数現れた。ODRの意義を理解していない学習者は、ODR多読へのコミットメントも低かった。

本実践では、Odyssey Reading Worksheetという多読用ワークシートを開発して、校種や学習者のレベルを問わず実践できる、生徒の多読意欲を喚起して英語力を向上させる多読授業の「標準化」を試みた。外部試験の数値や多読語数、アンケート結果から、この試みはある程度成功したといえる。

多様な背景の生徒が入学する公立高校において、トップダウン処理と統語処理を取り入れたワークシート型の多読活動が、読書意欲を維持しつつ英語力も向上させることができた。特に英語に苦手意識を持つ学習者が、長期間にわたって多量の多読を継続するためには、やさしい英語の絵本を買い揃えるといったハード面の整備だけでなく、多読の面白さや本の深い理解を体感させる指導のようなソフト面の充実が不可欠である。その点、読書意欲を刺激して4技能統合型の多読を展開できるODRは学習者にとって多読学習支援ツールとなり得るだろう。

8.2 今後の研究課題

今後の研究課題として、まず、ODRの長期的な検証が挙げられる。本実践のODRの開発と導入時期は、対象生徒が高校3年生に進級した後だった。彼らからのアンケートには「読むだけの通常の多読に物足りなさを感じていたので、4技能統合のODRは楽しかった。」「1年次から取り組みたかった。」という感想も出ており、生徒自身がODRによる多読学習の変化を実感していた。もしODRを高校1年次から3年間実践した場合、多読学習者にどのような変容が見られるのかといった長期的な検証は今後の課題であろう。

また、多読授業を成功させるために、学習者の実態を踏まえ、長期的な視点に立ってODRを活用することが求められる。ODRはライティングやスピーチ活動を含んだワークシート型多読であるため、短期的には学習者の授業内読書時間を減少させる。そのためODRの意義や目的を繰り返し丁寧に伝えて、授業外での多読図書貸出しを推進するなどの指導が求められる。また、学習者の実態に合わせて、適切な実施回数を設定することも必要であろう。

本実践はリーディング力の伸長を検証したものであるが、今後はスピーチやライティング力のデータも収集し、ODRの4技能への効果を検証したい。

さらに、ODRは多読学習用に開発したワークシートであるが、より汎用性を高めることで多様な用途に活用できる可能性がある。例えば、新聞雑誌記事や論説文、入試読解問題などに取り組む際、ODRは学習者の意欲を刺激し、より深い読みを体感させることができると期待できる。加えてODRの内容を簡略化し、どのような学習者も楽しく多読学習できる、小学生や中学生用のワークシート開発も検討したい。

本実践を基盤に、多様な学習者の英語力と意欲向上に貢献するため、引き続き研究を継続していく。

謝辞

このような研究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、選考委員の皆様に厚く御礼を申し上げます。特に、助言者の長勝彦先生には、論文執筆について多大な支援を頂戴しました。本研究の相談に乗ってくださった渡邊真美子先生、Ms. Summer James、そして都立大田桜台高等学校の9期生の皆さんとの多読への真摯な取り組み無くして、このOdyssey Reading Worksheetは完成しませんでした。最後に、3学年の進路指導で多忙を極め、論文執筆が遅々として進まなかった私を常に励まし続けてくれた、妻と二人の息子に心から感謝いたします。

参考文献(*は引用文献)

- Aebersold, J. A. & Field, M. L. (1997) "From Reader to Reading Teacher: Issues and strategies for second language classrooms". Cambridge press 多田穂(訳)『読みの学習者から読みの教師へ』東京:英宝社.
- Bartlett, F.C.(1932). *Remembering: a study in experimental and social psychology*. Cambridge : Cambridge University Press.
- Bransford, J.D. & M.K. (1972). Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall. *Journal of Verbal Learning and Behavior* 11, 717-726.
- Brown, H., & Cambourne, B. (1990). Read and retell: A strategy for the whole language/natural learning classroom. Portsmouth, NH: Heinemann.
- Carrell, P.L. & Eisterhold, J.C. (1983). Schema theory and ESL Reading pedagogy. *TESOL Quarterly*. Vol.17 (4), 553-573.
- Carrell,P.L. (1984). Schema theory and ESL reading: classroom. *The Modern Language Journal*, 68(4), 333-343.
- Furukawa, A., Takase, A., & Nishiawa, H. (2009). A Successful ER Program for Japanese Students of All Ages. Paper presented at the 43th Annual TESOL convention and Exhibit in Denver, USA
- Goodman,K. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 6, 126-135.
- Gough, P. B. (1972). One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.),*Language by ear and by eye: The relationship between speech and reading*. Massachusetts Inst. of Technology P.
- Kintsch, W., & van Dijk,T.A. (1978). Toward a model of text comprehension and production. *Psychological review*, 85, 363-394.
- Morrow, L. M. (1985). Retelling stories: A strategy for improving young children's comprehension, concepts of story structure, and oral language complexity. *Elementary School Journal*, 85, 647-661.
- Rumelhart, D.E., & Ortony, A.(1977). The representation of knowledge in memory. *Schooling and the acquisition of knowledge*. Richard Anderson, Rand J. Spiro and William Montague (Eds.) 99-135. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). Computational models of cognition and perception: *Parallel distributed processing: Explorations in the microstructure of cognition*, 1&2. Cambridge, MA: MIT Press.
- Zamel,V. (1992).Writing one's way into reading. *TESOL Quarterly*, 26, 463-487.
- 卯城祐司(1993)." Semantic Mapping を用いた英文読解指導の可能性—郡部校生のスキーマをモデルとして—"『僻地教育研究』52:129-136.
- 卯城祐司(編著)(2009).『英語リーディングの科学 読めたつもりの謎を解く』.東京:研究社.
- 門田修平・野呂忠司(編) (2001).『英語リーディングの認知メカニズム』東京:くろしお出版.
- 北原延晃(2010).『英語授業の「幹」を作る本(下)』東京:ベネッセコーポレーション.
- 酒井邦秀・神田みなみ(2005).『教室で読む英語100万語 多読授業のすすめ』東京:大修館書店.
- 高瀬敦子(2009).『英語多読・多聴指導マニュアル』東京:大修館書店.
- 津田塾大学言語文化研究所読解研究グループ(2002).『英文読解のプロセスと指導』.東京:大修館書店.
- 天満美智子(1989).『英文読解のストラテジー』東京:大修館書店.
- 中野達也(2014).「多読は速読に有効か? —音読・心内音読・黙読の違いが読解速度と内容理解に与える影響—」. STEP BULLETIN, vol.26,100-117.
- 西澤一・米澤久美子・粟野真紀子(2019).『図書館多読のすすめ かた』東京:日本図書館協会
- 増見敦(2014).「中高一貫校での英語多読指導の科学 —統計・言語分析による、英語力、言語力、自己効力感の変化を予測する科学的指導モデルの提案—」. STEP BULLETIN, vol.26,83-99.
- 吉田寿夫(2006).「研究法についての学習と教育のあり方について思うこと、あれこれ」 吉田寿夫(編著).『心理学研究法の新しいかたち』東京:誠信書房.
- 古川昭夫(2010).『英語多読法』東京:小学館.
- 文部科学省(2018).『高等学校学習指導要領(平成30年告示)』

資料1：必読書リストと推奨読書リスト

必読書

<Step1>

ORT Stage 1-3 96 冊

推奨読書リスト

<Step2> (このレベル終了までに10万語・500冊を目標)

ORT Explorer Stage 1-3

ORT Stage 4-5

ORT Explorer Stage 4-5

Oxford Read and Discover Level 1~2

Oxford Reading Tree Traditional Tales Stage 1-4

ORT Floppy's phonics Stage 1-4

Oxford Project X Stage 1-4

ORT Floppy's phonics Stage 1-3

ORT Read At Home

Scholastic Readers

I Can Read

Penguin Young Readers

Read it yourself Level 1-4

My First Read Books Biscuit

Leslie Patricelli

Elephant and Piggie

Step into Reading Step 1

Read at Home Level1-3

Clifford

<Step3> (このレベル終了までに15万語・700冊を目標)

ORT Stage 5

(Stage 5 と Stage 6 はレベル差が大きいので、6に進む前に推奨読書リストを出来るだけ読むことを目標)

ORT Explorer Stage 6-7

Ready to Read Level 1

Scholastic Reader Level 2-4

Oxford Read and Discover Level 3-4

Step into Reading Step 2

Foundations Reading Library Level 1-4

Curious George

Pigeon

Winnie the witch

Fly Guy

Info Trail [beginner, emergent]

Fireflies Stage1-5

ORT Floppy's phonics Non-fiction

Rookie Read

Puffin Easy to Read Level1

資料1：必読書リストと推奨読書リスト**<Step4>**

(このレベル終了までに20万語・1,000冊を目標)

ORT Stage 6

(Stage5 と Stage6 はレベル差が大きいので、6に進む前に必読書を終えておく)

ORT Explorer Stage 6-7

Foundations Reading Library Stage 5-7

Step into Reading Step 3

Oxford Read and Discover Level 4-6

Page Turner Level 1-4

<Step5> (このレベル終了までに25万語・1,300冊を目標)

ORT Stage 7-8

ORT Explorer Stage 7-8

Step into Reading Step 4-5

Getting to know the world greatest artist

Mr. Men & Little Miss

Penguin Readers Level 1

Oxford bookworms

Page Turner Level 5-8

<Step6> (このレベル終了までに30万語を目標)

ORT stage 9

ORT Explorer Stage 9

Page Turner Level 9-12

Lunch Lady

Mercy Watson

Magic Tree house

Rainbow Magic

Nate the Great

Flat Stanley

Cam Jansen

Frann K. Stein

Junie B. Jones

Who was

資料2：Odyssey Reading Worksheetを使った多読授業手順

Odyssey Reading Worksheet 手順				
	手順	分	具体的な手順	目的
プレーリーディング 3-5分	選書	3	書架から本を選ぶ。3分で3冊程度を目安にする。	
	本の情報記入	0.5	ODRに日付、本のタイトル、総語数を記入する。	
	type of structure 選択		選んだ本のストラクチャータイプを、Narrative / Informative / Argumentative から選び、ボックスに <input checked="" type="checkbox"/> を入れる	ストラクチャータイプの分析
	セマンティック・マッピング	1	タイトルと挿絵をよく見て、本の中で使用されていると予想した単語を、1分間で書き出す。想起できた単語数を記録する。マインドマップ状に、スピードを重視して書くとよい。※ここで本の中身は見ない	背景知識の活性化
	内容推測	1	タイトルと挿絵から、本の内容を推測する。推測した内容を1文で書く。推測すること自体に意味があるので、正解にこだわらなくてよい。※5W1Hなど	仮説設定による読書への積極的姿勢形成
	Question 作成	1	本の内容についての質問を作成する。質問を作ることが難しい場合、サンプルの質問を参考にする。	読書の目的づくりによる動機付け
インリーディング	多読	10-40	多読3原則に沿って多読を楽しむ。授業時間に合わせて多読時間を調整する。	
ポストリーディング 5-8分	セマンティック・マッピングの確認 (ボキャブラリーチェック)	1	読書後、想起していた単語が本文で使われていた場合、赤ペンで○を囲む。また重要と思われる単語を赤ペンで追記する。マッピングに単語を追加し、本の内容を構造化する。	想起した単語と実際の単語の比較、文章構造のマッピングによるサマリ一作成準備
	質問への回答	0.5	自分で作成した質問に対する答えを書く。	情報の抽出
	関連付け	0.5	本と自分とのかかわりを書く。はじめは「こじつけ」でもよいので、小さな関連性を見つけるとよい。	自己関連性の向上に動機付け
	評価	0.5	誰にお薦めしたいかという視点で評価する。	内容の客観的な認識促進
	コピーイング for Favorite sentence	0.5	お気に入りの1文を書き写す。ライティングや会話で使ってみたいという視点から探すとよい。	正しい文のコピーイング（書写）、センテンス集の土台作り
	サマリーライティング	0.5	本を要約する。まずボックスの真ん中に、Main idea を書く。何の話かを一言で表す。	最重要的要素の抽出
		0.5	type of structure を選び、N or I/A のどれかに <input checked="" type="checkbox"/> する。	ストラクチャータイプの分析
		0.5	Narrative の場合、setting, plot, characters, solution を <input checked="" type="checkbox"/> する。 ※Informative/Argumentative の場合、I/A に <input checked="" type="checkbox"/> し、要約の視点を（ ）に記入する。グラフィック・オーガナイザーの表を参考にしてもよい。	代表的な4つの視点を使用したサマリ一作成の容易化
		3	ボックスごとの視点から、本文の要約を書く。単語やキーワードではなく、主語+述部（動詞と目的語や補語、修飾語）で書く。※ここまでで1冊読み終えた学習者は、別の本を読む。	統語処理の促進
スピーキング活動 ※授業終了の5分ほど前に教員が合図	2	ペアを作り、相手とWSを交換する。本の要約を1分間話す。生徒Bは相手の英語を聞き取りながら、ワークシートへ絵を描き、相手のサマリーをヴィジュアル化してあげる。AとBの役割を交換する。	インフォメーションギャップを利用した有意味のやり取り、統語処理の促進	
振り返り	1	ODRを使った多読活動を振り返る。気づきを日本語で書く。 ※多読多聴手帳に読んだ本の語数とコメントを記入することも忘れない。	メタ認知の強化	

資料3：アンケート質問項目**アンケート質問項目**

1. 多読(Reading)は楽しい。
2. 多聴 (Listening) は楽しい。
3. 多読用の本を読む前に、絵やタイトルから内容を推測している。
4. 多読の時、「速読」を意識して読んでいる。
5. 多読の時、知っている単語の意味をつなぎ合わせて、なんとなく読んでいる。
6. 多読の時、英文を読んでいても理解できず、内容が頭に入っていかないことがある。
7. 多読の時、細かい所にこだわらず、大意(本の大まかな内容)をつかもうとしている。
8. 多読の時、意味の分からない単語は、一つ一つの意味を考えたりせず、飛ばして読んでいる。
9. 多読の時、意味の分からない単語は、前後の文やイラストなどから想像して読んでいる。
10. 多読の時、英文の主語や動詞など、文の構造を意識して読んでいる。
11. 多読の時、英文を日本語に訳して読んでいる。

資料4：ODR多読についてのアンケート結果

No.	質問項目	思全 わくそ なそ いう	思ま わあ なそ いう	ふ つ う	思ま うそ う	そ非 常 にう	小 計	平均 値
Q26	ODRのポストリーディング活動で、本の内容理解が深まる。	0	0	10	30	22	62	4.19
Q1	普通の多読に比べて、ODRを使って多読すると本の内容理解が深まる。	0	0	6	39	17	62	4.18
Q7	ODRで多読多聴を行うと、普通の多読に比べて Writingが伸びると思う。	0	4	12	17	29	62	4.15
Q22	セマンティック・マッピングの時、タイトルや表紙の絵をよく観察するようになった。	3	1	10	20	28	62	4.11
Q27	ODRでペアと Speaking活動をすると、本の理解が増す。	2	0	11	29	20	62	4.05
Q17	プレリーディング活動(ODRの表面)をすると、読む本の理解が深まる。	0	0	15	31	16	62	4.02
Q20	ODRのプレリーディングを行うと、頭の中で本の内容をイメージしやすくなる。	2	3	6	37	14	62	3.94
Q32	ODRを使って多読すると、普通の多読よりも本の内容を覚えていることがある。	1	4	10	30	17	62	3.94
Q5	ODRで多読多聴を行うと、普通の多読に比べて Readingが伸びると思う。	1	0	17	29	15	62	3.92
Q28	ODRでペアとアウトプット活動をすると、多読で読んだ表現が定着すると感じる。	2	4	14	23	19	62	3.85
Q8	ODRで多読多聴を行うと、普通の多読に比べて Speakingが伸びると思う。	0	6	15	24	17	62	3.84
Q19	プレリーディングで本文内容の予測をすると、内容理解がしやすい。	3	2	16	24	17	62	3.81
Q4	ODRを使うと、毎回の多読授業で達成感を得られる。	2	6	17	27	10	62	3.60
Q13	ODRを使わないで多読するとき、タイトルや表紙をよく観察するようになった。	3	10	14	20	15	62	3.55
Q18	本のストラクチャータイプ(N, I, A)を分析すると、本の理解が深まる気がする。	3	7	17	24	11	62	3.53
Q21	ODRのセマンティック・マッピング(単語の書き出し)すると、多読中に新しい単語を覚えやすいと思う。	3	9	19	15	16	62	3.52
Q33	ODRでペアとアウトプット活動をするほうが、普通の多読より楽しい。	4	9	17	15	17	62	3.52
Q15	プレリーディング活動(ODRの表面)をすると、多読の本が読みやすくなる。	3	6	21	21	11	62	3.50
Q23	セマンティック・マッピングすると、読む意欲が高まる。	2	6	24	20	10	62	3.48
Q11	ODRを使った多読を行い、4技能のうちに Writingに自信が持てるようになった。	2	9	22	17	12	62	3.45
Q9	ODRを使った多読を行い、4技能のうちに Readingに自信が持てるようになった。	2	5	26	22	7	62	3.44
Q16	プレリーディング活動(ODRの表面)をすると、多読の本を読む意欲が高まる。	4	6	17	29	6	62	3.44
Q2	普通の多読よりも、ODRを使った多読は楽しい。	0	13	20	19	10	62	3.42
Q12	ODRを使った多読を行い、4技能のうちに Speakingに自信が持てるようになった。	1	12	19	23	7	62	3.37
Q6	ODRで多読多聴を行うと、普通の多読に比べて Listeningが伸びると思う。	3	10	25	14	10	62	3.29
Q25	ODRを使うと、多読中に意味の分からぬ單語に出会っても、推測しやすくなった。	3	11	23	20	5	62	3.21
Q14	ODRを使わないで多読するとき、読む前に内容を推測したり質問を考えたりするようになった。	5	13	17	19	8	62	3.19
Q10	ODRを使った多読を行い、4技能のうちに Listeningに自信が持てるようになった。	4	9	28	15	6	62	3.16
Q30	ODRのアウトプット活動後、「もっと多読をしたい」と感じる。	5	13	21	16	7	62	3.11
Q3	多読する意欲が乗らない時も、ODRを使うと読書する意欲が出る。	6	9	25	17	5	62	3.10
Q29	ODRのWriting活動、Speaking活動を行うと、多読の時に英文の主語・動詞などを意識するようになる。	13	9	14	16	10	62	3.02
Q34	高校卒業後も多読多聴を続けるつもりだ。	10	12	17	19	4	62	2.92
Q24	ODRを使うと、多読中に速読を意識するようになった。	8	15	19	16	4	62	2.89
Q35	もしODRをつかうことができれば、今後も一人で多読学習を続けることができると思う。	7	21	16	14	4	62	2.79
Q31	Post readingの「自分とのかかわり」を考えると、本文の理解が深まる。	11	19	18	11	3	62	2.61

第32回 研究助成

B 実践部門・報告V・英語能力向上をめざす教育実践

高等学校での内容言語統合型学習(CLIL)による 他教科と連携した授業の実践と効果の検証

研究者:栃木県／栃木県立 宇都宮南高等学校 教諭 渡邊 聰代

《研究助言者:和泉 伸一》

概要

本研究では、新学習指導要領を踏まえ、公立高等学校における英語の授業改善を目指し、他教科と連携した授業実践の過程とその効果を明らかにした。具体的に授業では、他教科の持つ内容と英語の学びを統合し思考と協学を促す、内容言語統合型学習(CLIL)により、本研究者が担当するコミュニケーション英語Ⅰの授業を現代社会および保健体育の担当教員と共に実践した。授業はコミュニケーション英語Ⅰの検定教科書を用い、現代社会との連携で2回、保健体育との連携で1回行った。それらについて①授業準備の記録、②授業観察、③生徒への質問紙調査、④連携した他教科教員へのインタビューおよび本研究者による授業のふりかえり、を記録した。その結果CLILによる授業での他教科と連携した教科横断的な学びが生徒の学習動機を高め、学びが教科間でつながり、広がりのあるものとなるということが分かった。また、今後のさらなる実践と検証の必要性が挙げられた。

1

はじめに

高等学校において2022年度より全面実施される新学習指導要領では、「主体的・対話的で深い学び」を実現させるための授業改善が求められている(文部科学省, 2018a)。これを受けて外国語科では各科目の内容構成が一新され、「主体的・

対話的で深い学びの実現に向けた授業改善を推進するため、学習過程全体を通して、知識・技能を実際のコミュニケーションにおいて活用し、思考・判断・表現することを繰り返し経るような指導の改善・充実が図られる必要がある」(文部科学省, 2018b, p.9)と提示されている。

内容言語統合型学習(Content and Language Integrated Learning, 以下 CLIL)は、この「主体的・対話的で深い学び」を可能にする教育的アプローチであり、近年日本の高等学校でも実践が徐々に増えつつある。「CLILとは、ヨーロッパを起源とする外国語指導法で、言語学習と教科内容を統合させ、そこに思考活動と協学、異文化理解を取り入れ、学習者の体験的学習の促進を目的の1つとしている」(山野, 2013, p.94)。英語を使い、必然性のある多様で豊かな内容について思考しながら協働で学ぶことにより、生徒の学習動機や学習効果を高め、思考・判断・表現に必要な知識・技能としての英語の習得を目指すことができる。さらに、新学習指導要領では「コミュニケーションを図る資質・能力を育成する」という目標達成のために、指導計画作成上の配慮事項として「言語活動で扱う題材は、生徒の興味・関心に合ったものとし、国語科や地理歴史科、理科など、他の教科等で学習した内容と関連付けるなどして、英語を用いて課題解決を図る力を育成する工夫をする」(文部科学省, 2018b, p.129)とされている。吉田(2017)も、これからの日本の英語教育の方向性としてCLILの重要性を挙げ「CLILなどの考え方を基に、他教科と連動した学習内容について、

英語という道具を使って、思考、判断、そして発表することが求められる」(p.128)としている。

しかし、栃木県の公立高等学校の英語の教員は、英語の文法や語彙について熟知してはいるものの、他教科の内容について広く深く精通している訳ではない。一方、他教科の教員は、普段の授業では教科の内容を日本語で教えており、それらを英語で教える技能は持ち合っていない場合が多い。そこで、具体的にどのようにすれば、生徒の発達段階や知的好奇心に応じた、他教科等での学びを高等学校の英語の授業に入れ、目標達成を目指すことができるのかについては、英語の教員が他教科の教員と連携して模索すると効果的であると考えた。これに基づいて本研究では、公立高等学校の「コミュニケーション英語Ⅰ」の授業において、CLILにより、他教科と連携した授業を実践し、実践の過程とその効果を明らかにする。コミュニケーション英語Ⅰの検定教科書で扱っている内容を他教科の担当教員と共有し、関連する内容を扱っている部分の授業計画について助言を受けたり、実際に授業でチームティーチングを行ったりすることが、授業改善につながるかを調べる。とりわけ、生徒の興味・関心を高め、教科横断的な知識の習得と深い学びに結び付き、授業が生徒にとって「教科等間での学びのつながりや広がりがあるもの」(文部科学省, 2018b, p.129)となるかを検証する。

2 研究の背景

2.1 CLILとは

CLILは欧洲連合(以下EU)が1990年代半ばに、複言語主義、平和社会の実現を目指す言語教育政策を推進する中で発展してきた外国語教授法であり、現在EUではほとんどの加盟国の言語教育で採用されている(杉谷, 2010; 笹島, 2011)。第2言語習得理論からもその有用性が期待できる(和泉, 2016a)。池田(2016)によるとCLILは、教室での学習活動を通して知識力・思考力・語学力・協働力等の学力を養い、21世紀を生き抜く地球市民として必要な、認知的能力・社会的能力・倫

理的能力等の「21世紀型汎用能力」(池田, 2016, p.15表2参照)を育むことを目指している。これは、新学習指導要領が目指す「知識・技能」「思考力・判断力・表現力」「学びに向かう力・人間性」の育成も包括しているであろう。

CLILの授業では、理科や社会など他教科の題材や時事問題、国際理解など意味のある本物の内容と、習得をねらう第2言語を、4技能を通して使いながら学ぶ。内容と言語を統合し、教科教育と言語教育の両方を実現させる。加えてCLILでは、後述する「4つのC」というフレームワークを用いて計画的に思考活動と協学を入れる点が大きな特徴であり、これにより授業で英語を本物で豊かな内容を通して学び、意味ある活動の中で思考しながら活用することが可能となる。授業で体験的活動を通して、英語の「知識力」とともに思考力と表現力などの「活用力」を伸ばすことができるのである。

2.2 CLILの「4つのC」

CLILでは、「4つのC」と呼ばれる、Content(内容), Communication(言語), Cognition(思考), Community(協学)のフレームワークに沿って、指導案を作る。池田(2011, p.5)も、CLILは4つのCを「有機的に結び付けてパッケージングした点」が画期的であり、「この枠組みに即して教材を作り、授業案を考え、指導を行えば、相乗効果により高品質の教育が実現されるというのがCLILの要諦である」と述べている。池田(2011)によると、4つのCはそれぞれ次のようなものとされる。

- ・Content(内容)=学習者の興味と動機付けを高める本物の内容
- ・Communication(言語)=言語の「学び(知識の習得)」と「使う(活用)」の融合
- ・Cognition(思考)=記憶だけでなく、分析や創造などさまざまな思考を促すタスク
- ・Community(協学)=オーセンティックな教材を用いた4技能統合型の協同学習

和泉(2016b, p.76)は次のように述べている。「語るべき内容(Content)があるからこそ、他者

に伝える手段としての言葉(Communication)が必要になる。また、考えるべき内容があるからこそ、思考(Cognition)が刺激され、広がりと深まりを増していく。さらに、他者との交わり(Community)があるからこそ、コミュニケーションが重要になっていくのである。要するに、4つのCは、それぞれが支え合って存在しているが故に、教育の質の向上を掲げるCLILで大事な指針となるのである。」従って、「4つのC」を適切に理解し、統合して授業に入れることができ、CLILの授業を実現させる上で必要不可欠となる。

2.3 高等学校のCLIL実践

池田(2011, p.10)によるとCLILには「目的」「頻度・回数」「比率」「使用言語」によるバリエーションがあり、図1のように示すことができる。

Soft CLIL	【目的】	Hard CLIL
英語教育		科目教育
Light CLIL	【頻度・回数】	Heavy CLIL
単発的／少数回		定期的／多数回
Partial CLIL	【比率】	Total CLIL
授業の一部		授業の全部
Bilingual CLIL	【使用言語】	Monolingual CLIL
英語・日本語		英語

■図1:CLILのバリエーション (池田,2011, p.10)

池田(2013)はまた、現時点では英語の教員が他教科の内容を教える形態(Soft CLIL)が日本の高校の教育現場には導入しやすいとしており、埼玉県立和光国際高等学校の英語教育においてSoft CLIL実践の指導助言(山崎, 2013)を行っている。

日本の高等学校で採用するコミュニケーション英語の検定教科書は、題材やトピックは他教科の内容となっているので、すでにCLILで授業実践が可能な状態である。例えば、2019年度、本研究者の所属する栃木県立宇都宮南高等学校の

コミュニケーション英語Iで採用している検定教科書 *Vivid English Communication I* (第一学習社)でも、スポーツ、歴史、日本文化、科学、健康、音楽、食物、環境問題、災害、平和など、さまざまなテーマが選定されており、理科や地歴公民科、家庭科、芸術科、保健体育科などの各教科で教えられている内容となっている。これらを「英語について」学ぶ訳読や分析的学習に使うのではなく、CLILで豊かに、体験的に学ぶようにさせたい。そして、その学びを「学びに向かう力・人間性の醸成」につなげたい。しかしながら「ことばの教師」である英語の教員にとって、主体的・対話的で深い学びを促すために踏まえる必要のある「内容」については、知識が十分でなく、自分でそれらを習得する時間も足りないことが日常的な悩みである。この点で、 笥島(2011)は埼玉県立川越女子高等学校においてCLILにより英語教師と生物の教師がチームティーチングを行った事例を紹介しているが、生物の教師が英語教師と共に授業を実践することによりCLILの授業を実現させている。また、チームティーチングも形態は多様で、必ずしも同じ授業を協働で実施するという必要は無く、事前の授業準備で連携するなどのやり方もあると述べている。こういった教師間のさまざまな協力形態を探ることもCLIL授業を実現する上で大切な試みとなる。

3 研究の目的とリサーチクエスチョン

本研究の目的は、栃木県の公立高等学校の英語教師がCLILを用いて、他教科の教師と連携した授業をどのように実践することができるか、そしてそれは生徒や教師にとってどのような効果があるかを明らかにすることである。この目的を達成するために、以下のリサーチクエスチョンを設定し、それらに沿って実践の過程の記録とアンケート調査等の結果を示すこととした。リサーチクエスチョンは以下の通りである。

1) 授業の前で

RQ1

他教科と連携した授業を準備するにあたり、授業の内容はどのように設定できるか

RQ2

授業の準備において他教科教員との連携はどのように進めるべきか

RQ3

内容と言語を統合するために、指導案はどのように作成するべきか

RQ4

役割分担はどのようにするべきか

2) 授業で

RQ5

指導案にどのくらい沿った授業ができ、どのような課題がでてくるか

RQ6

授業に対して生徒たちはどう反応するか

3) 授業の後で

RQ7

生徒たちは授業をどのように感じ、何を学んだのか

RQ8

担当教員2人は授業についてどのような感想を持つのか

4

研究方法

4.1 協力者

本研究は、栃木県立宇都宮南高等学校1学年生徒(1クラス39~40名)と、英語以外の教科を担当している教員2名の協力のもと、行われた。教員2名はそれぞれ地歴公民科、保健体育科の担当教員である。

4.2 研究授業の手順

研究授業は現代社会との連携で2019年10月31日と2019年11月27日、保健体育との連携で2019年12月12日に行われた。はじめは、本校のコミュニケーション英語Iで採用している検定教科書 *Vivid English Communication I* (第一学習社)で扱われている題材を基に本研究者が他教科の教員に協力を依頼した。そして、協力に同意を得

てから話し合いながら授業準備をすすめ、研究授業を実践した。研究授業の直後に授業を受けた生徒に質問紙による調査、連携した他教科教員にはインタビューによる調査を行った。

4.3 データ収集法と分析法

本研究では、①授業準備の記録、②授業観察、③生徒への質問紙調査、④連携した他教科教員へのインタビューおよび本研究者による授業のふりかえりを行い、その結果を検証した。授業準備の記録は、授業準備の打ち合わせの際に本研究者が記録したノートや交換したメモ、授業で使用するため準備したパワーポイントを用いたスライドに基づいて記録した。これによって、RQ1~RQ4で問う、他教科と連携した授業の準備における授業内容の設定、他教科教員との連携、指導案作成、そして教員間の役割分担について明らかにしようとした。授業観察は1台のビデオカメラにより記録されたデータにより行われたが、RQ5とRQ6で問う、実際の授業の進行と課題、及び生徒の反応を明らかにしようとした。生徒への質問紙調査では、授業に対する生徒の感想と意見を5段階のリッカート尺度を使った選択回答式10項目と自由記述回答式4項目を通して尋ねた。選択回答式アンケートで実際に使われた質問は後述の5.2で示す表1~3にある通りである。選択回答式アンケートの分析を通して、RQ6にある授業に対しての生徒たちの反応、記述式アンケートへの回答をもとに、RQ7の生徒たちの学びと意見について明らかにしようとした。研究授業後に行なった他教科教員に対するインタビューと本研究者のふりかえりをもとに、RQ8にある担当教員の授業についての感想を明らかにしようとした。

5

結果

5.1 研究授業

実施した3回の研究授業について授業の準備と実践した授業の実際について、経過に沿って説明する。研究授業は準備段階から日本語は使わ

ないことを前提にし、基本すべて英語を用いて行ったが、授業中に生徒が分からぬであろうと予測された単語が出てきた場合や指示の意味を正確に伝えたい場合、教材として用いた動画に日本語が入っていた場合、予定していなかったことを本研究者以外の教師が伝えることになった場合は、日本語の使用があった。

5.1.1 現代社会と連携した授業

現代社会と連携した授業は、2回行った。1回目は2019年10月31日に行い、そこでの反省を踏まえて、同じ内容で発問を改善した2回目の授業を11月27日に実施した。これらの授業について、授業の準備、1回目の授業、2回目の授業の順に述べる。

連携したいと考えた現代社会担当のI教諭は、本校の学習指導部長でもあり、教科横断的な学びを推奨している。そのため研究授業を実施する以前に、教科書のLesson 3: Look Cool on Your Streetsを扱う際に、ファッショニから児童労働を考える授業をI教諭と連携して実践したことがあった。今回、教科書の題材の中からI教諭と連携することで合意した題材は、Lesson 8: The Sphinx in Dangerであった。この課では、生活排水の増加やダムの建設により塩害が広がり、そのために崩れつつあるエジプトのスフィンクスが題材として取り上げられている。前回ファッショニをテーマにした授業を実施した際は、本研究者は全て英語を使って授業をしたが、I教諭は

授業の後半部分は日本語で授業を行った。しかし、その授業の前半の様子から、準備をしていればI教諭の英語の発話が生徒に十分伝わると分かったので、今回の授業はI教諭に全て英語を使って授業をしていただくこととした。それはI教諭には大きな挑戦だった様子である。

授業はLesson 8を扱う最初の授業であり、大半の生徒はまだ教科書を読んでいないことを前提に計画した。I教諭と授業準備のための打ち合わせは全部で主に5回行った。1回目は、I教諭に教科書Lesson 8の内容を伝え、話し合いにより授業で扱うテーマを「開発と環境」ということに決めた。2回目の打ち合わせでは、「開発と環境」を考えるためにイースター島の話が興味深いので扱いたいとの提案がI教諭からあった。そこで、次の打ち合わせまでに授業で使うパワーポイントによるスライドを本研究者が作成し、I教諭に見ていただくこととした。3回目はそのスライドを見てもらい、提示の順番を修正し、生徒への問い合わせをするか話し合った。また、イースター島の話はI教諭が紙芝居を作成し、伝えたいことを本研究者が訳し、I教諭がそれを音読して生徒に聞かせる形態を取ることとした。I教諭が発話して生徒が分かるように英語ができるだけやさしいものを使うよう心掛けた。4回目はスライドを確認しながら再度生徒への問い合わせを2人で考えた。最終的に、What can we do to make this world better?で終わることとした。授業前日となる5

Content (内容)	Communication (言語)	Cognition (思考)	Community /Culture (協学／文化)
(宣言的知識) エジプトのスフィンクスやイースター島の歴史について知り、開発と環境破壊について知識を得る。	(言語知識) ・語彙(世界遺産) ・表現(関係副詞where)	(低次思考力) 出来事やスピーチを理解する。	(協同学習) ・ペア ・グループ ・一斉
(手続的知識) 開発の必要性と環境保全について、自分がやるべきことを考える。	(言語技能) ・聞く ・読む ・話す ・書く	(高次思考力) これまでの環境破壊を分析する。 自分の考えについて論じる。	(国際理解) 世界の開発や環境問題に目を向ける。

■図2: CLILの授業設計(現代社会との連携:1回目)

回目には、指導案と授業手順を確認した。

授業1回目は、2019年10月31日(木)5時間目(13:40-14:20)に、第1学年、出席者39人の学級で行われた。また、本研究の指導助言者である上智大学和泉伸一教授と、本研究者がCLILについて普段指導助言いただいている宇都宮大学山野有紀准教授及び山野研究室ゼミ生に授業観察をしていただく形で行われた。池田(2016, p.17)のフレームワークによるCLILの授業設計図は図2のとおりである。

授業では、I教諭の紹介後、前の週にI教諭が修学旅行へ行ったことを伝え、Kyoto is the city where people can enjoy many beautiful temples and shrines.と、関係副詞を使った文を導入した。その後、スライドに沿って本研究者がDo you like traveling? Where would you like to visit? Is it a World Heritage Site?と生徒に問いかけた上で、I教諭はイースター島に行きたいこと、本研究者はエジプトに行きたいことを場所の写真と共に示した。エジプトではスフィンクスのすぐ近くにファストフード店がある様子や、スフィンクスはその上にライオンの像があったかもしれないという説があることなどを紹介した。次に、台風で栃木県が被害を受けた様子や、太陽光パネルで鉄砲水が出るようになってしまった場所があることを写真と共に紹介し、Do you think it is OK?と生徒たちに問いかけた。それから、I教諭が紙芝居でイースター島の歴史を紹介した。緑豊かであったイー

スター島が、人間が原因で荒廃していく様子と、それは現代の地球全体にも起こり得ることかもしだれないということをI教諭が英語で説明した。本研究者は生徒の様子を見ながら、難しいと思われるところは説明を加えたり、くりかえし説明したりした。これにより、「開発と環境保全」が共存しにくい現実を理解した上で、最後にWhat can we do to make this world better?という問い合わせについて生徒たちはグループで考え、発表した。“We can reduce plastics and do volunteer activities.”や“We must make situation environment-friendly.”のような意見が生徒たちから引き出せた。このグループワークで50分の授業は終了した。

授業後、授業観察者から助言・感想等をいただいたが、その中で特に、授業に多くを盛り込み過ぎて、話の流れが分かりづらく、授業最後の発問も大きく漠然としたものになってしまって、全体としてまとまりのない授業になってしまっていることが指摘された。さまざまな内容を授業で扱いたいと欲張ったことが裏目に出た形となってしまった。そこで再度、スライドや発問を吟味して授業計画を練り直し、2回目の授業を1回目実施の学級とは別の学級で実施した。

授業2回目は2019年11月27日(水)5時間目(13:40-14:20)に、第1学年、出席者40名の学級で行われた。CLILの授業設計は図3のとおりで、下線部が1回目の反省から改善したところで

Content (内容)	Communication (言語)	Cognition (思考)	Community /Culture (協学／文化)
(宣言的知識) エジプトのスフィンクスやイースター島の歴史について <u>推測しながら知識を得る。</u>	(言語知識) ・語彙(世界遺産) ・表現(関係副詞where)	(低次思考力) 出来事やスピーチを理解する。	(協同学習) ・ペア ・グループ ・一斉
(手続的知識) <u>歴史についての知識を生かし、これからの社会はどうあるべきか考える。</u>	(言語技能) ・聞く ・読む ・話す ・書く	(高次思考力) 歴史上の出来事を分析する。 自分の考えについて論じる。	(国際理解) 世界の開発や環境問題に目を向ける。

■図3: CLILの授業設計(現代社会との連携:2回目)

ある。

前半の手順は1回目と同様に行きたい場所を話題に関係副詞を導入し、I教諭がイースター島に、本研究者がエジプトに行きたいことを伝えた。その後は台風や太陽光パネルの話題は省き、代わりにエジプトの歴史を示し、今は砂漠でも昔は緑が広がっていたことを紹介した。そしてその後、現在のイースター島の写真を使い、モアイ像のみが残され、人間が誰も住んでいないことに触れ、What do you think had happened on this island?と問い合わせ、グループで考えて意見をまとめ、発表させた。生徒たちからは“People cut a lot of trees to carry Moai. So the island's nature was broken.”や“Rats ate a lot of plants. It broke nature. People brought rats by ship.”のような意見が出された。その後、I教諭がイースター島の歴史についての紙芝居を紹介し、最後の場面に移る前に、What do you think will come next?と発問した。紙芝居の最後の場面は、現代の都市部でビルが乱立する様子が描かれており、これがモアイ像の残されていたイースター島の様子と重なると言っている人がいるが、そう思うか問い合わせる場面となっている。その後、I教諭が紙芝居の次の場面に何を描いたかをグループで考えさせ、発表の時間を持とうとしたが、授業終了時間が迫っていたので、発表は省き、紙芝居の最後の場面を紹介して終わる形となった。

5.1.2 保健体育と連携した授業

Lesson 1: You Can Do Something for the Olympicsについては、生徒たちは本校に入学してすぐの4月に学んだが、保健体育科のT教諭と連携したいと考え、復習・まとめという位置づけで12月にもう一度扱うこととした。T教諭はバドミントン競技の国際審判の資格があり、世界中で活躍しており、審判として必要な英語力を持っている。また、ロンドンオリンピックのバドミントン競技で起きた、無気力試合の審判の経験を持っているので、そのことについて教えていただき、その是非について生徒に考えさせることでCLILの授業を実現できると考えた。そこでまず保健体育の教科書を借りて、オリンピックの扱われている部分について本研究者が学んだ。保健体育の教科書では、オリンピックムーブメントやオリンピックの起源、現在のオリンピックの在り方などが扱われていた。英語Iの教科書では2020年に東京オリンピックが開催されることや、古代オリンピックと現代オリンピックの違い、英語を学んでボランティアをしよう、といった内容が扱われているのだが、それよりは保健体育の教科書の内容を中心に授業で扱った方が生徒の興味関心を高められると考えた。T教諭の多忙なことを考えると、じっくり打ち合わせの時間は持てないことが予測されたので、3回に分けて必要なことを教えてもらったり、パワーポイントのスライドに必要な資料を提供していただいたりすることにした。1

Content (内容)	Communication (言語)	Cognition (思考)	Community /Culture (協学／文化)
(宣言的知識) オリンピックの歴史や種目、オリンピズムについて知識を得る。	(言語知識) ・語彙(スポーツ) ・表現(助動詞can)	(低次思考力) 説明やスピーチを理解する。	(協同学習) ・ペア ・グループ ・一斉
(手続的知識) 2020年の東京オリンピックに向けて、自分がやるべきことを考える。	(言語技能) ・聞く ・読む ・話す ・書く	(高次思考力) これまでのオリンピックの歴史を分析する。 自分の考えについて論じる。	(国際理解) オリンピックムーブメントや世界平和に目を向ける。

■図4: CLILの授業設計(保健体育との連携)

回目は3分ほど職員室で立ち話をするような様態であったが、保健体育の教科書でオリンピックについて扱われている内容について伺った。クーベルタン男爵が平和を理念の1つに掲げ、古代オリンピックを復活させたこと、しかし、現在に至るまで2度の大戦によってオリンピックが中止になってしまったこと、現在でもオリンピックにはお金がかかり、国の経済事情に大きく左右されている側面があること、そしてその解決策の1つとして、世界のスポーツエリートを育てるユースオリンピックがあること、ユースオリンピックでは若者に対してスポーツだけでなく倫理的な教育も行うことを教えてもらった。また、T教諭がロンドンオリンピックで撮影した写真や、無気力試合を扱ったニュース番組の動画のデータを提供してもらった。2回目は本研究者が1回目の打ち合わせを基にパワーポイントを用いて授業用スライドを作成し、体育教官室でT教諭に見てもらい、内容について話し合った。T教諭はそのスライドを持ち帰り、生徒に伝える内容を自ら英語で入力した。3回目は授業直前に授業で用いる機材について簡単に調整した。

授業は2019年12月12日(木)6時間目(14:40-15:20)に、第1学年、出席者37名の学級で行われた。CLILの授業設計は図4のとおりである。

授業の前半は全て本研究者が行った。その間、T教諭は教室の後ろから生徒たちの様子を観察していた。授業の最初では本研究者がDo you like sports? What sport do you like?と問いただし、生徒とやりとりをした後、自分が好きなスポーツがオリンピック種目になっているか確認してみよう、として、オリンピック競技種目を確認した。その後、人々は何故スポーツをするのか、オリンピックは何のためにあるのかを問い合わせ、スポーツには世界と未来を変える力があるということをおさえた。さらに、教科書で扱った古代と近代オリンピックについて簡単に触れ、オリンピズムとオリンピックムーブメント、2つの大戦下ではオリンピックが中止になったことについてスライドを使いやり取りしながら確認した。ここで、T教諭に登場していただいた。ロンドンオリンピックのバドミントン審判団の写真をスライドで示し、T教諭がWhere am I?と生徒に問い合わせ、皆で写真の中にT教

諭の顔を見つけた。ロンドンオリンピックで開会式のチケットが20万円もしたこと、T教諭は審判なのでそれを支払わずに開会式を観ることができたことを紹介した後、バドミントンの試合の写真を見せた。Something very sad was happening. とT教諭が述べた後、無気力試合についてのニュース動画を見せた。この試合によって中国と韓国の選手は失格となり、それによって日本が銀メダルを取れたことも合わせて伝えた。この後、T教諭からWhy do you play sports? Why do you watch sports?という問い合わせがあった。T教諭は続けてIt is very important for players to win. But thanks to the support of many fans, players can play. Fans can get energy, joy and excitement from players. Badminton, the Olympics are not just for players.と述べた。そしてこの後、生徒たちはグループでWhat can we do for Olympics?という問い合わせに対してグループで考え、発表した。生徒たちからは“We can learn about the Olympics.”や“We can cheer, support, and believe athletes.”のような意見がでた。最後に時間が3分程余ったので、T教諭からユースオリンピックの理念について日本語で講義をしていただき、授業を終えた。

5.2 生徒による質問紙回答結果

研究授業後にアンケートを実施し、以下のような結果が得られた。

5.2.1 選択回答式アンケート結果

選択回答式アンケートの結果は、表1~3の通りである。

■表1: 選択回答式アンケート結果(現代社会:1回目)

		平均	標準偏差
①	授業の内容に興味が持てた。	4.36	.66
②	授業の内容が分かった。	4.33	.69
③	先生の言っている英語が分かった。	3.87	.79
④	今までに学んだ英語が役に立った。	4.10	.63
⑤	内容について考え、自分の意見を持つことができた。	4.08	.62
⑥	伝えたいことを英語で何と言ったら良いか、考えることができた。	3.79	.85
⑦	伝えたいことがあったが、英語にすることができなかつた。	3.59	1.05
⑧	自分の意見などを英語で何と言えばよいかさらに知りたくなつた。	3.97	.82
⑨	クラスの一員として、この授業に参加することができた。	4.32	.62
⑩	クラスメイトと協力してグループの意見をまとめることができた。	4.30	.73

■表2: 選択回答式アンケート結果(現代社会:2回目)

		平均	標準偏差
①	授業の内容に興味が持てた。	4.73	.63
②	授業の内容が分かった。	4.65	.53
③	先生の言っている英語が分かった。	4.30	.51
④	今までに学んだ英語が役に立った。	4.25	.58
⑤	内容について考え、自分の意見を持つことができた。	4.40	.66
⑥	伝えたいことを英語で何と言ったら良いか、考えることができた。	4.28	.63
⑦	伝えたいことがあったが、英語にすることできなかつた。	3.13	1.08
⑧	自分の意見などを英語で何と言えばよいかさらに知りたくなつた。	4.28	.55
⑨	クラスの一員として、この授業に参加することができた。	4.48	.55
⑩	クラスメイトと協力してグループの意見をまとめることができた。	4.40	.73

■表3: 選択回答式アンケート結果(保健体育)

		平均	標準偏差
①	授業の内容に興味が持てた。	4.38	.63
②	授業の内容が分かった。	4.41	.59
③	先生の言っている英語が分かった。	4.08	.67
④	今までに学んだ英語が役に立った。	4.14	.58
⑤	内容について考え、自分の意見を持つことができた。	4.11	.80
⑥	伝えたいことを英語で何と言ったら良いか、考えることができた。	4.16	.72
⑦	伝えたいことがあったが、英語にすることできなかつた。	3.46	.95
⑧	自分の意見などを英語で何と言えばよいかさらに知りたくなつた。	4.00	.81
⑨	クラスの一員として、この授業に参加することができた。	4.11	.69
⑩	クラスメイトと協力してグループの意見をまとめることができた。	4.22	.84

5.2.2 記述回答式アンケート結果

次に示すのは、記述回答式アンケートの結果である。記述回答式アンケートでは、

- ①この授業で扱った内容について、自分が学んだと思うこと、または考えたことは何ですか。
- ②この授業で学んだ英語はどんなものですか。(カタカナでも良いので書いてください。)
- ③現代社会(保健体育)の内容を英語で学ぶことについて、どう思いますか。
- ④この授業の感想を書いてください。(先生たちへのアドバイスもお願いします!)

の4つの質問に回答を求めた。

現代社会と連携した授業: 1回目

①の質問に対しては、「自分が思っているより環境問題が深刻化していること、解決するためには多くの時間と人の力が必要であること」「これから私たちが生きる世界の問題だから、しっかり考えないといけないと思った」など、環境問題について、または自分にできることについて記述したものが32回答あった。②の質問に対しては、protect, environment, emissionなど、内容を表現するために使った語句を挙げたものが32回答と大半を占めた。文を挙げたものは3回答で、「We can ~」「We should protect forest」「We can clean オール over the world.」であった。③の質問に対しては、「日本語では難しいと感じることが英語ではわかりやすく感じた」「今まで習った英単語がたくさん出てきたので社会の授業を英語で受けるのはとても楽しかったです」「一石二鳥だと思いました」「英語を理解しようとするからとても印象に残った」「あまり現社は好きで無くとも興味が持てた」など、良いと評価するものが31回答あった。一方で、「理解できないところもあったので日本語の方がよいと思った」「難しい単語が多い」「分からぬ單語が多く、調べなければならないので少し大変だと思った」など、難しさを挙げたものが5回答あった。④の質問に対しては、良かったとする回答が32回答得られた。理由としては、「英語で現代社会の授業をやるという斬新な考えに驚いた」「1回の授業で2つの教科を学習することは、難しいが興味が湧く」など、他教科との連携について挙げているものが10回答、「絵とか写真とかが多くてとても分かりやすかったです！楽しかったです！」「紙芝居がとても面白く、わかりやすかった」など、教材の工夫について挙げているものが9回答、「自分なりに世界のためになることをしっかり考えることができたのでよかった」「自分の意見を考えることができてとても楽しかった」など、思考が良かったとするものが7回答あった。

現代社会と連携した授業: 2回目

①の質問に対しては、「今の社会もイースター島のモアイの話と同じで、将来、建物だけが残っ

て人間がいなくなるのではないかという話で自分もそうかもしれないなと感じた」「イースター島の歴史の凄絶さに驚かされた」など、具体的な内容や自分の意見についての回答が37回答あった。また、協学を通して他の生徒やグループの意見が聞けて良かったとする記述も3回答見られた。②の質問に対しては、deforestation, disaster, sinkなど、内容を表現するために使った語句を挙げたものが34回答であった。③の質問に対しては、「自分で考えたことを英語にする力がついて、とてもおもしろいと思った」「普段学んでいる英語を応用して使うことができるのでいいと思う」「いつもの授業よりも深く考えることができた。自分の考えをもつことができた」「興味が持てた。言葉が分からぬ分、ちゃんと聞こうと思った！」など、良いと評価するものが37回答あった。一方で、難しいというものが3回答あったが、「英語で学んでいるのでわかりにくくなつた分、理解しようとしたと思う。」といった回答も3あった。最後に、④の質問について、良かったとする回答が37回答得られ、その理由として「モアイの謎について考えることができた。もっと深く知りたいなと思った。」など、思考が良かったとするものが9回答あり、「みんなと考えて、発表したり、それを聞いたりすることはとてもおもしろかった」など、協学が良かったというのも9回答あった。さらに、「2つの教科が組み合わさったことで、余計に難しくなるかと思ったが、この方が理解しやすかった」など、教科横断の試みについて挙げているものが7回答、「先生の絵やテレビの画面もあり、とても楽しく分かりやすかった」など、教材の工夫について挙げているものが4回答あった。

保健体育と連携した授業

①の質問に対して、「スポーツは世界や人を変える力があること」「選手の無気力試合は他の人の不幸にもなってしまうということ。スポーツマンシップは大事ということ」「オリンピックがただのスポーツ大会ではないこと」など、内容について、または内容を踏まえて自分で考えたことを述べた回答を36回答得た。②の質問に対しては、athlete, cheer, Youth Olympicsなど、内容を表現するために使った語句を挙げたものが36回答

と大半を占めた。③の質問に対しても、「英語を学ぶと同時にオリンピックについて考えることができ、一石二鳥で良い時間になった」「保健体育は好きなので、またやりたい」「英語でこのような内容を学ぶことは多少なりとも将来につながると思う」など、良いと評価するものが37回答あった。一方、「頭が疲れる」「難しい単語をたくさん覚えないといけないと思う」など、難しいとしたものが3回答あったが、「身近なことを英語で表すのは難しいけどすごく楽しかった」と、難しいが良いとしたものが1回答あった。④の質問に対しては、良かったとする回答が33回答得られた。その理由として、「英語×保健という2つの教科がとても合っていて、面白かった」「1つの科目を集中してやることも必要だが、この先の将来に向けて2科目をつなげて行っていくことも大切だと思う」など、教科横断の試みについて挙げているものが11回答、「テレビや写真映像を使った説明で分かりやすかったし楽しかった」など、教材の工夫について挙げているものが5回答、「いつもとは違う、考える時間の多い授業で楽しかった。」など、思考が良かったとするものが4回答、「スポーツについての内容だったので、興味をもって聞くことができた」など、内容についてのものが6回答、「テレビで映像を見たり話し合いしたりするのが良かった」と、教材に加え協学が良かったというものが1回答あった。また、特徴的だったこととして、「Lesson 1でやったような内容をより深く学ぶことができ、ためになった」「Lesson 1で習った事を生かせて授業をうけることができて良かった」と、既習事項の応用に意義を見出している回答が2回答あった。

5.3 指導者のふりかえり

授業後のインタビューでは、現代社会で連携したT教諭からは示唆に富んだ感想をいただいた。普段の現代社会の授業では、教師が説明をし過ぎてしまったり、引き出したい答えに誘導し過ぎてしまったりすることがあるようである。一方で英語科と連携した授業では、英語が堪能では無いからこそ英語教師の力を借りながら何とかしてゆっくり粘り強く伝えようとしたり、生徒が表現しようとするのを待ったりできたということであ

ある。また、日本語ではなかなか自分の意見が言えなかった生徒が、英語で発言している様子が観察できるなど、英語の持つ特性が効果を発揮しているように思えたそうである。加えて、まずはやってみないと分からないので、やってみることが大切で、やった後で改善点や以後の対策を話し合うことが授業力の向上につながるということを実感できたということであった。さらに、今後はこのように単発ではなく、他の科目とも体系的に連携を考え、教科間での学びが生徒にとってもっとダイナミックにつながりや広がりのあるものとなるよう工夫していくべきだと考えているということであった。

保健体育で連携したT教諭からも、前向きな感想をいただいた。T教諭にとってはスポーツの楽しさや素晴らしさを生徒に伝えることが最も大切なことなので、その手段が英語でも日本語でも、実際に競技を体験することでも、どれでも自分にとって同様に達成できることが重要であるということであった。また、今回英語科と連携したことにより、スライドや動画を通して英語で段階を踏んで分かりやすく自分の体験を伝え、生徒に深く考えさせる機会を持てたことは喜ばしいことであったということである。

本研究者が他教科と連携した授業をCLILで実践して最も意義があると感じたことは、言語の教師でありながらも、内容を生徒に伝え、考えさせたいという自身のモチベーションが高まったことである。これまで、イースター島やオリンピックについてあまり深く知らず、生徒に伝えたいと思うほどの情熱が湧かずにいた。しかし、現代社会と保健体育の教師から伝えたい、考えさせたいという熱意を強く感じ、歴史の現代社会とのつながりやオリンピックの奥深さなど、興味深い内容を具体的に提示していただいたので、本研究者の授業準備と授業を実践する意欲が高まった。また、授業中は内容を伝えることに意識が向くので、教師も生徒も自分が発話する英語の正確さにあまりこだわらず、伝えるためにたくさん話すことができた。これは、訳読みや文法等についての分析的学習を行っている授業では実現が難しいと思われる。一方で課題に感じたことは、言語指導の内容の中への組み入れである。内容についての話が進み、生徒の発話が促されている時に授業の流

れを損なってしまう恐れがあったため、言語についての指導やフィードバックの時間を十分に取ることができなかった。英語の授業で他教科と連携したので、50分の授業の中での時間配分や、年間や単元ごとの指導計画の中に言語指導を上手く組み入れることを周到に準備したい。

6 考察

結果を踏まえて、RQ1からRQ8について考察する。

RQ1:
他教科と連携した授業を準備するにあたり、
授業の内容はどのように設定できるか

研究授業はどの授業もコミュニケーション英語Ⅰの授業を行った。従って、授業の内容は、検定教科書で扱っている内容と関連したものを使うことを前提とした。

現代社会と連携した授業では、本研究者がコミュニケーション英語Ⅰの検定教科書をI教諭に手渡し、世界史を専門とするI教諭から、Lesson 8: The Sphinx in Dangerで連携したいという要望をいただいた。この結果から言えることは、他教科の教員にまずは英語の教科書を見てもらい、連携しやすい内容を見つけていただくことで授業の内容設定をすることができるということである。

一方で、保健体育と連携した授業では、T教諭が以前に行ったロンドンオリンピックについての日本語での講演会を本研究者が聞いていたので、Lesson 1: You Can Do Something for the Olympicsの内容を扱うことに決めてからT教諭に連携をお願いした。英語担当の教員が扱う内容をあらかじめ決めてから、他教科の教員にその内容をお願いするという形でも内容設定をすることができる。さらに、保健体育の教科書を借りてオリンピックについて本研究者が学んだことにより、英語の教科書と共通する内容や関連性を授業で扱うことができたので、他教科の教員に扱う内容について、他教科で使用しているテキストを紹介していただき、英語の教員が英語の検定教

科書と他教科のテキストで関連しているところを見つけ、授業で扱う内容として設定することもできると言えるであろう。

RQ2:

授業の準備において他教科教員との連携は
どのように進めるべきか

本校の教員はいつもとても忙しい。普段担当している授業に加えて本研究者と連携した授業をしていただくので、大きな負担にならないよう注意して授業の準備を進めた。また、他教科教員の英語発話力に応じて授業準備を進めることを約束して授業準備を進めることで、英語の授業に教員として参加する心理的な負担を軽減するよう努めた。

現代社会と連携した授業の準備は、I教諭が教科横断的な学びを深める必要性を感じていたこともあり、本研究者からの打ち合わせの要望に積極的に応じてくださり、授業を改善して同じ内容を扱って2回目の授業も実践できた。1回目の授業の前にも十分に打ち合わせをしたつもりだったが、足りなかった。この結果から、可能な限り打ち合わせは十分に行い、他教科教員との連携を進めていくべきだということが言える。

保健体育と連携した授業の準備では、T教諭がロンドンオリンピックで撮った写真や、試合についてのニュース動画など、オーセンティックな素材が豊富にあった。その中から授業の流れにあったものを本研究者が選び、授業で使うスライドに教材として組み込んでいった。また、T教諭は多忙な方であったが、授業の合間などは快く本研究者の質問や打ち合わせの要請に応じて下さった。この結果から言えることは、授業でやりたい内容がはっきり決まっていて、英語の教員がそれを把握している場合は、密に打ち合わせをしながら授業を作り上げる形態ではなく、英語の教員の方がまず計画を練って、分からぬ細かいところを他教科の教員から助言をいただきながら進めていくという方法を取ることもできるということである。

RQ3:
内容と言語を統合するために、指導案はどうのように作成するべきか

豊かな内容についての学習を英語学習と結び付けるために、それぞれの授業において指導案作成の段階で Genesee & Hamayan (2016, p.166) に倣い、内容についての目標は現代社会および保健体育の教員が、言語についての目標は英語の教員である本研究者が設定した。

現代社会と連携した授業では、本課の言語材料である関係副詞は、行った場所や行ってみたい場所について関係副詞 where を用いた文にして、スライドで 3 例示し、板書もして導入した。また、内容豊かな英語を生徒に理解させるために、紙芝居をするというアイディアは I 教諭が出した。I 教諭が文の原案と絵の部分をつくり、本研究者が英訳して制作した。さらに、2 回目の授業では問い合わせの工夫を行い、生徒の思考を深める流れを作ることができた。この結果から、他教科の持つ豊かな内容を英語教育に活用するには、英語教員と他教科の教員でアイディアを出し合い、それを英語の教員が CLIL の「4 つの C」を意識して精査すべきだと言える。また、個々の活動が独立してしまい授業の流れが滞ってしまわないよう、活動を授業に「編み込む (interwoven) (Coyle, Hood, & Marsh, 2010)」ことが重要であると考える。

保健体育と連携した授業は、4 月に学んだ Lesson 1 のまとめの授業として 12 月に行ったこともあり、Lesson 1 全体で使用されている助動詞 can の定着を目標としたが、生徒は概ねできるようになっていた。生徒への問い合わせが What can we do for the Olympics? であったこともあり、生徒の表現の中に助動詞 can があることが多かった。この結果から、定着を促したい文法事項や言語材料は、難易度が低い場合や定着を促したい場合は、生徒が注目して使用することになるので、問い合わせの中に組み込むと有効であると考える。また、ロンドンオリンピックの無気力試合のニュースについての動画は日本語のものだったので、生徒はその後 T 教諭が使用した disqualified や participate などの語彙の意味を予測することが

できていた。この結果から、生徒の英語での理解を助けるために、効率的な日本語の使用も考えて指導案を作成すべきであるということが言えるであろう。

RQ4:
役割分担はどのようにするべきか

他教科の教員と連携して CLIL で授業を進めるにあたり、他教科の教員との打ち合わせで話し合って役割分担を決めたが、その際特に他教科の教員の英語運用に対する自信の度合いを配慮した。

現代社会と連携した授業では、I 教諭は英語の正確さには自信が無かったが、英語を授業中に発話することに対しては前向きであった。そのため、紙芝居に必要な英語を書いておいて、それを音読していただいた。この結果から、他教科の教員が、授業中英語を発話し、内容の教員として授業に関わりたい意欲がある場合は、事前に英語を用意し、それを授業で示す役割を担ってもらうことができると言える。

保健体育と連携した授業では、T 教諭は英語を発話することに慣れており、自分でスライドに発話する台詞を英語で入力して準備できていた。従って、はじめはスライドに沿って本研究者が発話して授業を進め、途中 T 教諭に話し始めていただく場面を決めておき、そこからは T 教諭に中心となって話していただいた。生徒も T 教諭が話す英語に感動していた。この結果から、他教科の教員が比較的英語を話せる場合は事前に役割分担を決めておき、英語の教員は他教科の教員が中心となって話す場面では生徒の理解を見取りながら、語彙の難易度によっては介入して説明を加えるなど、補助的な役割を果たすと良いということが言えるであろう。

RQ5:
指導案にどのくらい沿った授業ができ,
どのような課題がでてくるか

どの研究授業においても、授業のはじめに warm up として生徒に身近で簡単な質問をしてインタラクションを行った。その後、他教科の教員から教えていただいた内容を、主にパワーポイントのスライドを使ってインタラクションをしながら示していった。そしてどの授業でも必ず、CLIL の「4つの C」の Cognition(思考) を促す問いを生徒に投げかけた。また、授業の内容理解を促すために、現代社会と連携した授業では、紙芝居を使った。保健体育と連携した授業では、動画を使った。これらの教材はいずれも、生徒の思考を促し、協学で問い合わせの答えを考え出すために有用であった。この結果から、授業は事前に考えていた指導案に概ね沿った展開で進められたと言えるであろう。

課題となったことは、他教科の教師が話す英語をどのように英語教師が生徒に分かるように補助するかである。他教科の教師が話す英語が生徒に伝わっていないと思われる場合、英語教師が介入して難易度の低い表現に言い換えをしたり説明を加えたり、時には日本語で意味を言ったりすることが必要である。しかし、どのタイミングでどのくらいの時間を取りて介入するか、またどのようにしたら授業の流れを損なわずに介入できるか、今後の研究と熟練が必要である。

RQ6:
授業に対して生徒たちはどう反応するか

ビデオ録画の観察からは、どの授業でも、生徒たちは積極的に活動に参加していたことが分かる。教師の発話する英語については、特に他教科の教師の言うことを理解しようとする姿勢が見られた。また、問い合わせに対する答えをグループで考え、発表する活動では、生徒がそれぞれ考え、グループに何らかの貢献をしていた。このことから、CLIL による他教科と連携した

授業実践により、生徒の積極的な授業参加や学び合いを促進できると言うことができるであろう。

5.2.1 の選択回答式アンケート結果からは、生徒は概ね、授業の内容に興味が持て、内容が分かったようであると分かる。また、教員の発話する英語を理解し、今までに学んだ英語を応用する機会を得ていたと言えそうである。さらに、内容について考え、自分の意見を持つことができたかということについて、また、伝えたいことを英語で何と言ったら良いか考えることができたかについても、ほとんどの生徒ができていたようである。一方で、伝えたいことがあったが、英語にすることができなかった生徒は本研究者が予想していたほど多くなかった。しかし、自分の意見などを英語で何と言えばよいかさらに知りたくなったという生徒は多かったので、授業中は既存の知識で何とかコミュニケーションを図っていたということがうかがえる。クラスの一員として、この授業に参加することができたかということと、クラスメイトと協力してグループの意見をまとめることができたかということについても、よくできていたということが言えそうである。つまり、CLIL による他教科と連携した授業で「4つの C」がそれぞれ具現化され、それを通して生徒たちは教科横断的な学びを深めることができていたということが言えよう。

RQ7:
生徒たちは授業をどのように感じ,
何を学んだのか

記述回答式アンケート結果から、生徒たちは CLIL による他教科と連携した授業を良いと感じ、内容と言語の両方を学んだということが分かる。

授業の内容についての質問に対する結果からは、生徒たちが授業中、知識の確認や吸収をし、考えを深めたり自分の意見を持ったりすることができていたことが分かる。特に注目したいのは現代社会と連携した授業の1回目と2回目で生徒の記述回答の内容が異なっていることである。2つの授業で扱った題材は同じものであつ

たが、展開と発問が異なっていた。これを反映して、1回目では環境問題について、または自分にできることについて記述したものが多く、2回目では具体的な内容について、または自分の意見についての回答が多かった。本研究者が考えていた以上に、内容についての学びの深まりは、題材そのものだけでなく、授業の展開や教師の発問に左右されていた。

学んだ英語についての質問に対する結果からは、生徒が主に語彙を中心に行なったと感じていることが分かる。また、興味深いことに、本研究者が授業の学習課題として設定した文法事項よりも、教師が使ったり、生徒が自分の意見を表現するために使ったりした語句の方を圧倒的に生徒たちが「学んだ英語」として選んでいることが分かる。さらに、数は少ないが、グループ活動で個別に本研究者が質問を受けてフィードバックした表現や文法事項について回答しているものもあった。従って、CLILによる他教科と連携した授業では、教師が生徒の学びを意図した文法事項よりも、生徒が自ら注目した語句や、生徒が自分で疑問に感じた表現や文法事項の方が記憶に残りやすいということが言えると考えられる。

他教科と連携した授業についての質問に対する結果からは、大半の生徒が良いと考えていたことが分かる。ただし、生徒たちの中には、内容を完全に理解したいという欲求があり、英語が難しいためにそれが阻害されることへの不安があると考えられる。その解消のために、多少のあいまいさには慣れさせるということに加え、教材の工夫や、教師が使用する英語の難易度の調整、効果的な日本語の使用など、生徒の理解への足場かけを増やす必要性もあると思われる。

授業の感想についての質問に対する結果からは、ほとんどの生徒たちが、教科横断ということや教師による教材の工夫によって、内容と英語を結び付け、協学を通して思考することを求められていたことについて、良いととらえていたことが分かる。また、「難しいけど楽しかった」という意見が複数あったことから、生徒にとって難しい内容や英語が動機の低下に結びつくとは一概に言えないということが分かった。むしろ達成可能な範囲で「チャレンジングな(和泉、

2009)」内容や英語を用意することは有用であると言えるであろう。

RQ8:

担当教員2人は授業についてどのような感想を持つのか

指導者のふりかえりから、教師が成果であると考えていることは2点挙げられる。1つはCLILにより内容の教師と英語教師が連携することは可能であり、それによって、生徒の興味・関心が高まり、単独で教えるよりも明らかに積極的に生徒が授業に参加しているなど、相乗効果が得られるということである。教師が伝えようとする英語も、日本語に比べ難しいからこそ、生徒たちが集中して聞き入っている様子が観察できた。もう1点は教科間の連携により、教材の工夫や授業の手法を学び合えるので、教師としての資質や授業実践のモチベーションを高められるということである。どの教科でも、またどの教師も、それに教材の工夫や協学のやり方を研究してきている。しかし、今までそれらを共有する場が少なかった。今回実際に授業で共有したおかげで、お互いに学び合ったことがあり、今後それを別の授業でも応用することが可能であろう。

一方で、今後継続して実践していくにあたり課題となることとして考えられることが2点ある。カリキュラムの中での位置づけと、教員の意識である。他教科の教師はそれぞれ、自分の教科の持ち時間に加えて、授業時数が増えた形で英語との連携授業に応じてくださった。これは今回が単発の授業であるので実現したことで、これを継続して実践するためには、学校のカリキュラムの中に位置づけ、組織全体で行う必要があるだろう。また、教員が教科間連携に意義を見出しているか、その熱意はどのくらいか、ということによって教材研究などの授業準備の質と量が変わってくるだろう。まずは多くの先生方に、CLILにより連携することによって授業が豊かになり、ひいては生徒の学びがつながり広がりを持つようになるということを観察や体験によって実感していただくことが第一歩となると考える。

7

まとめ

本研究ではCLILにより英語の教師が他教科の教師と連携した授業を試み、その実践の過程と生徒及び教師への効果を検証した。その結果、CLILにより他教科の内容と英語を統合し、思考を刺激しながら協学により授業を展開したことによって、生徒にとって授業での学びが教科間でつながり、広がりのあるものとなるということが分かった。また、教師にとっても、お互いを刺激しあい、授業力を高めることができる試みとなることが分かった。

栃木県の公立高等学校における英語授業では、本研究者の授業も含め、未だ語彙や文法等についての分析的説明や和訳に多くの時間が費やされ、生徒の体験的学習が不足している。また、体験的学習を中心とした授業においても、その体験の質が教育的営みとして十分に生徒の興味関心に合った、課題解決力を育むものとなっているかについては疑問が残る。それには教師の多忙なことや研修機会の不足など様々な理由があり、すぐに改善することは難しいかもしれない。しかしながら本研究を通して、実は自分の周囲に豊かな内容を持った他教科の教師が存在し、教師同士も学びあうことによって、目の前の生徒たちに合った興味深い授業を実践し、学びを深めさせることができることが分かった。特に、本研究者は、これまであまり知らなかった、連携した他教科教師の自分の教科に対する熱意に刺激を受け、力を借りることで自分の授業力が思いがけず向上するという経験をすることができた。

このような教科横断的な学びを継続的に実現させていくためには、本研究で行った以外の教科とも連携し効果を検証することや、学校全体での取り組みを推進することなどが必要である。新学習指導要領で求められる「主体的・対話的で深い学び」実現の一助となるよう、今後もさらなる実践を続けていきたい。

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、特に助言者となりご指導くださった和泉伸一先生に、篤く御礼申し上げます。和泉先生からの貴重な助言無しに本研究は遂行し得ませんでした。また、宇都宮南高校の先生方と生徒の皆様には本研究にご協力いただき、ここに心より感謝申し上げます。特に磯寿人先生、遠井努先生には共に授業者として多大なるご支援をいただきました。最後に、CLILによる授業実践について長年ご指導くださっている宇都宮大学の山野有紀先生にこの場を借りて深く感謝申し上げます。

参考文献(*は引用文献)

- * Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.
- * Genesee, F., & Hamayan, E. (2016). *CLIL in context: Practical guidance for educators*. Cambridge: Cambridge University Press.
- * 池田真(2011).「CLILの基本原理」渡部良典・池田真・和泉伸一(共著)『CLIL(内容言語統合型学習)』上智大学外国語教育の新たなる挑戦—第1巻 原理と方法』1-13, 東京:上智大学出版。
- * 池田真(2013).「CLILの原理と指導法」「英語教育」第62巻第3号, 12-14 東京:大修館書店。
- * 池田真(2016).「CLIL活用の新コンセプトと新ツール」池田真・渡部良典・和泉伸一(編)『CLIL(内容言語統合型学習)』上智大学外国語教育の新たなる挑戦—第3巻 授業と教材』1-29東京:上智大学出版。
- * 和泉伸一(2009).『「フォーカス・オン・フォーム」を取り入れた新しい英語教育』東京:大修館書店。
- * 和泉伸一(2016a).『第2言語習得と母語教育から「言葉の学び」を考える』東京:アルク。
- * 和泉伸一(2016b).『フォーカス・オン・フォームとCLILの英語授業~生徒の主体性を伸ばす授業の提案~』東京:アルク。
- * 文部科学省(2018a).『高等学校学習指導要領(平成30年告示)』
- * 文部科学省(2018b).『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』東京:開隆堂。
- * 杉谷眞佐子(2010).『EUの言語教育政策』関連事項の解説』大谷康照・杉谷眞佐子・脇田博文・橋内 武・林桂子・三好康子(編)『EUの言語教育政策』1-8東京:くろしお出版。
- * 笹島茂(編)(2011).『CLIL 新しい発想の授業』東京:三修社。
- * 築道和明・宇都宮正朗・馬本勉・五井千穂・河野省治・立川三・鳥生典江・藤本幸伸・宮本隆昌・森眞二・Joseph Lauer(2017).『Vivid English Communication I』東京:第一学習社
- * 山野有紀(2013).「小学校外国語活動における内容言語統合型学習(CLIL)の実践と可能性」『EIKEN BULLETIN』vol.25, 94-126 東京:公益財団法人日本英語検定協会
- * 山崎勝(2013).「CLILを使った「異文化理解」の授業実践と課題」『英語教育』第62巻第3号, 30-31 東京:大修館書店。
- * 吉田研作(2017).「これから日本の英語教育の方向性」坂本光代・渡部良典(編)『応用言語学から英語教育へ 上智大学英語教授法TESOLコースの過去・現在・未来』122-135 東京:上智大学出版

C 調査部門・報告 I • 英語教育関連の調査・アンケートの実施と分析

学習者の視点から見た英語による 専門科目(EMI)に必要な英語力のニーズ分析

研究代表者:東京都／早稲田大学大学院 在籍 太原 達朗

共同研究者:埼玉県／早稲田大学本庄高等学院 非常勤講師 工藤 秀平

東京都／早稲田大学大学院 在籍・日本学術振興会特別研究員 守屋 亮

《研究助言者:和田 稔》

概要

本研究の目的は学習者の視点から見た日本における英語による専門科目 (English-Medium Instruction, EMI) に必要な英語力を把握することであり、そのためにニーズ分析 (Needs Analysis) を実施した。研究対象は EMI プログラムを開講する学部・研究科を持つ関東の私立大学で、参加者は EMI の実施形態のうち、ほとんどの科目を英語で開講する English-Taught Program (ETP) 型と、選択必修科目に EMI が多い Semi-Structured (SS) のいずれかのプログラムを経験した学生もしくは卒業・修了生であった。本研究は複数の段階におけるデータ収集・分析を行った。最初にインタビューデータから EMI で経験する 47 個のタスク一覧を作成し、それを元に各タスクにおける頻度・難度を 4 段階評価で回答するアンケートを作成し実施した。各タスクを頻度・難度の平均値の高低を元に 4 種類に分類し、ETP 型と SS 型の両 EMI 実施形態における共通点・相違点を明らかにした。同時に、自由回答記述形式で EMI に必要な能力・必要なサポートの意見も調査した。本研究は日本の大学において強化の優先順位が高い EMI に必要な英語力を特定し、同時に、今後のニーズ分析の方法論的貢献も示唆した。

1 はじめに

今日、日本の大学において英語による専門科目 (English-Medium Instruction, 以下 EMI) が増加している。EMI とは「英語が第1言語ではない国や地域において、アカデミックな専門科目を英語で学ぶこと」(Dearden, 2014, p. 2) と定義される。日本でもグローバル 30 (文部科学省, 2011) やスーパーグローバル大学創成支援 (文部科学省, 2014) などの留学生の誘致・グローバル人材の育成を推進する政策が制定された結果、EMI を開講する大学のプログラムが増えてきた (Morizumi, 2015; Ota, 2018)。日本の大学における EMI の研究は、EMI の実施状況や政策研究などの広い視点から見たマクロレベルの研究 (例: Hashimoto, 2013; Rose & McKinley, 2018; Shimauchi, 2018) と、反対に、ある EMI 環境における特定のトピックに焦点を置いたミクロレベルの研究 (例: 動機づけ [Kojima & Yashima, 2017], スピーキング不安 [Kudo, Harada, Eguchi, Moriya, & Suzuki, 2017], 語彙能力 [Uchihara & Harada, 2018], 語用論的能力 [Taguchi, 2014]) が実施されてきた。しかし、このマクロレベル・ミクロレベルの中間に位置し、両者の研究の橋渡しを行う研究は少ない。特に、実際の教材・カリキュラム開発や EMI 受講者への取り組みへ

と繋がる、日本の大学における「EMIで必要な英語力」とは何かを調査する実証的な研究はほとんどない。日本のEMIにおいて学習者が事前にどのような英語力を身につけている必要があるか、またEMI受講中にどのような英語力に到達している必要があるかといった情報は、今後のEMIプログラムの運営・実施に示唆を与えると考えられる。これらを踏まえて、本研究は日本におけるEMIに必要な英語力を調査するためにニーズ分析(Needs Analysis)を行うことを目的とする。

2 先行研究

2.1 EMI

本節においてEMIの説明を行う。最初に、EMIの定義を確認する。EMIの定義は国・地域・文脈によって変わる(Macaro, Curle, Pun, An, & Dearden, 2018)。EMIの定義の違いを理解するためにはEMIに近い概念と比較対照するとわかりやすくなるため、その隣接領域の概念も同時にレビューを行う。

2.1.1 EMIの定義

EMIでよく引用される定義は「英語が第1言語ではない国や地域において、アカデミックな専門科目を英語で学ぶこと」(Dearden, 2014, p. 2)である。文字通りの解釈であればEMIはあくまで専門科目の内容を学ぶことを目的としており、言語を学ぶことは対象外である。しかし、一部の研究者や各地のEMI実践者は専門科目の内容と共に言語を学ぶことも範囲に入れて考えている。

EMIに関連する、言語と教科内容を共に学ぶアプローチとしてはContent-Based Instruction(CBI、内容重視指導)とContent and Language Integrated Learning(CLIL、内容言語統合学習)がある。CBIとCLILはいずれも言語と教科内容を同時に学ぶアプローチであることは共通している(例:Coyle, Hood, & Marsh, 2010; Snow, 2014)。しかし、CBIとCLILの由来は異なり、CBIは北米における教育的ニーズを満たすため

に発生したのに対し、CLILはEUによる複言語・複文化主義の理念を達成するための政策から生まれた(Brinton & Snow, 2017)。

このような外国語教育のアプローチに対して、EMIもCBIやCLILと同様に言語も学習の目的としていると主張する研究者もいる。CBIとEMIの関係については、例えば、Brinton and Snow (2017)はEMIはCBIの一種であり、豊かな学術的言語と複雑な概念を学ぶことは外国語学習に必要な条件だと述べている。CLILとEMIの関係についても、言語を学ぶこともEMIの目的だと主張する研究者がいる。具体的には“CLIL-ised EMI”(Moncada-Comas & Block, 2019)、つまりCLIL寄りのEMIという指導法も提唱されている。Airey (2016)は学習内容と言語は密接に絡み合っており、EMIを学習内容だけに焦点を当てるのは人為的・理論的な分類だと述べている。このように研究者によってはEMIにおいて専門内容を通して英語を学ぶことを副産物と捉えている(Murata, 2019)。日本においても、学習者に対するアンケートでEMIの受講目的は言語を学ぶことが一番多い回答だったという報告があった(Chapple, 2015)。

EMIの目的を専門科目のみを学ぶことと捉えるか、それとも言語と専門科目を両方学ぶことと捉えるかでは、EMIにおける英語力の位置づけは変化する。前者なら必要な英語力はEMI受講前に身につけておく必要があり、後者ならEMIプログラムを通して英語力を身につけるということになる。このように、EMIの目的をDearden (2014)の定義通りに専門科目のみを学ぶことか、それともCBI・CLILと近い概念かを意識するのは、EMIプログラムを実施する際に「EMIに必要な英語力」を決定するのに重要である。

また、EMIに近い概念で英語を学ぶことを重視するものがあり、例えば、English for Academic Purposes(EAP)と「オールイングリッシュ」がある。EAPは「学術目的のための英語」で、「オールイングリッシュ」は日本における「英語の授業は英語で」に対応する。EAPは特定の目的ための英語(English for Specific Purpose, ESP)の一部で学術英語に特化したものを目指す(Brown, 2016)。一方、オールイングリッシュは「教員は教室内で行う授業運営、授業内容の説明、

または生徒との会話を全て英語で行うことが基本理念である。また、生徒も英語を使うことが基本である」(藏本, 2017, p. 48)という方針の実践形態である。ただこれらは言語を学ぶことが主目的であり、専門内容を学ぶことはあまり意識していない。

しかし、注意すべきは、いずれの授業形態でも専門科目と英語の学習が完全には分かれてはいないことである。EMIがCBIとCLILのように言語が学びの対象に入るか断言できないのと同様に、EAPやオールイングリッシュを含む、普段の英語授業において専門内容を学ぶことが完全に省かれてきたわけではない。そういう意味では、言語と専門内容がどれだけ授業内容に入るかは程度の問題であると言える。

2.1.2 日本におけるEMIの実施形態

EMIは定義だけではなく、その実施形態も複数の種類が存在する。日本におけるEMIの分

類はBrown and Iyobe (2014)によると“1. Ad hoc”, “2. Semi-structured”, “3. Integrated”, “4. + a program”, “5. English-taught program (ETP)”, “6. Campus-wide”的6種類に分類される(表1)。同じ大学であっても学部や研究科が異なるとEMIの運営方法は変わることがある。そこで、各EMIの種類に応じた問題に対応する必要が出てくる。世界的にEMIは“4. + a program”, “5. English-taught program (ETP)”, “6. Campus-wide”的3種類を指すことが多いが、日本においてはより英語使用の割合が少ない“1. Ad hoc”, “2. Semi-structured”, “3. Integrated”的3種類も多い(Hino, 2017)。そのため英語の使用の割合が少ない種類のEMIにおける研究も必要である。特に、EMIの最終的な受益者である学生の英語にふれる量や授業形態などによって必要な英語力も変わってくる。この点から、EMIの実施形態による違いを意識したEMIのニーズ分析が必要であると言える。

■表1: 日本におけるEMIの実施形態の6分類

カテゴリー	内容	例
1. Ad hoc	授業を担当した教師が英語で教えることを選択した授業。 カリキュラムとの関連は薄い。	多くの大学に存在する
2. Semi-structured	専門科目に関する複数の授業が英語で開講される。専門内容の専門家(日本人・外国人)か語学教師(外国人)が授業を行う。 主に選択授業だが、必修になることもある。	早稲田大学教育学部 英語英文学科
3. Integrated	学内にある所定のEMIプログラム。多くの英語授業は選択制だが、入学・卒業時におけるベンチマークは存在する。修了証にもEMIであることが記される。	各大学の短期プログラム等
4. +a program	学部・学科間における英語による交換プログラム。 留学生用のコースも開講され、修了証が発行される。	大阪大学大学院工学研究科 フロンティアバイオテクノロジー国際プログラム
5. English-taught Program (ETP)	学位を得るためにほぼ全ての学科の授業が英語で与えられる。	早稲田大学国際教養学部、 上智大学国際教養学部
6. Campus-wide	キャンパスのほぼ全ての学部の授業が英語で教えられる。	国際教養大学

(注1)Brown and Iyobe (2014), Hino (2017)を元に作成。(注2)EMIの形態は必ずしもこの6つに分類しきれないこともある。例えば、早稲田大学教育学部英語英文学科はEMIの選択必修を多く受講する必要がない分野もある。一方、EMIであることは修了証には書かれないので、英語力のベンチマークは測定しており、英文成績証明書はEMIで受講した科目はEMIだと明記される。

2.1.3 日本におけるEMIの先行研究

日本におけるEMIの研究はマクロレベル(社会・政策レベル)とミクロ(実地の特定のトピックレベル)の調査が行われている。日本におけるEMIの導入が促進された契機はグローバル30(文部科学省, 2011)とスーパーグローバル大学創成支援(文部科学省, 2014)の2つの政策である(Morizumi, 2015; Ota, 2018)。マクロレベルの研究はこのような文部科学省が公開した政策文書を詳細に分析にした研究(Hashimoto, 2013; Rose & McKinley, 2018)や、これらの文書から読み取れる国際化戦略を決める外的・内的要因の分析(Shimauchi, 2018)などである。また、EMIの実施における課題は「1. 言語の課題」「2. 文化的課題」「3. 運営の課題」「4. 施設面」の4種類に分類される(Bradford, 2015, 2016; Brown, 2018)。これも広い目で見たEMIのマクロレベルの研究だと言える。それに対してミクロレベルの調査はEMI環境における特定の要因の関係や変化に着目した実証研究である。具体的には、EMI環境における語彙能力(Uchihara & Harada, 2018)、語用論的能力(Taguchi, 2014)、動機づけ(Kojima & Yashima, 2017)、不安(Kudo, et al., 2017)などのテーマに対して研究がなされてきた。

しかし、マクロレベル・ミクロレベルの研究を橋渡しする、教育機関レベルにおけるEMIの実施報告は日本において不足している。数少ない報告例として、第一にHarada(2017)は2016年4月に開始された早稲田大学教育学部英語英文学科の大規模なカリキュラム改訂の際に新たに設計されたCBI・EMIプログラムのカリキュラムデザインの方法と課題を詳細に述べた。もう一つの例では、Poole(2017)が同志社大学の例を挙げて、Global 30を始めとするマクロレベルでのEMI推進の政策と、各大学におけるミクロレベルでのEMI実施の方針に乖離があるため、高等教育機関でのEMI実施に問題が生じていると論じている。だがこれらの報告例はEMIの実施の初期段階であり、具体的にどのような能力が必要かという点はまだあまり明らかにされていない。特に学習者の視点からの、具体的にどのような英語力を運用するかという分析は日本においてまだ不足している。日本におけるEMI実施の問題点として、学生の英語力の低さによる

ドロップアウトが挙げられている(Hino, 2017; Tsuneyoshi, 2005)。その一方で実証的な研究が不足していることを考慮すると、学生を視点とした必要な英語力を把握する研究を行うのは喫緊の課題である。そのためにはどういった英語力が必要なのか、同時に異なるEMIの実施形態の中で、入学前までの英語学習歴や入学目的・学習目的、EMIに対する授業が異なる学習者が英語を使い専門科目を学ぶ上でどういった点に困難を覚えるかを調査するニーズ分析が必要である。

EMIのニーズ分析に関する研究ではEMIを実施する側(大学レベル・機関レベル・教師レベル)と受講する側(学生レベル・留学生レベル)の2種類に分かれる。日本における報告例としてSawaki(2017)は英語入試で問うべき英語力を定義するため、日本における私立大学の教員6名を対象として、理系分野で専門を学ぶのに必要な英語力をインタビュー調査した。結果、4技能を使うと答えた教員もいれば、リーディングが中心になると回答した教員もいた。また、いくつかの報告例においては学習者を対象としているが、学生がEMIに関してどのように受け止めているかが中心であり(例:Chapple, 2015; 村田, 飯野, & 小中原, 2017)、どのような英語力が必要かという観点からはほとんど研究が少ないので重要な点である。学生を対象に具体的なスキルを調査したニーズ分析の例はSuzuki, Harada, Eguchi, Kudo, and Moriya(2018)がある。この研究は私立大学におけるEMI授業を受講する大学生15名に自由記述式アンケートを実施し、EMIにおける苦手意識のあるスキルを明らかにした。実際のカリキュラム開発・教材開発等に繋げるためには学生のスキルベースの主観的なニーズに加え、後述するタスクベースの客観的なニーズを調査する必要がある。

2.2 ニーズ分析

2.2.1 ニーズ分析の定義

ニーズ分析の定義は複数存在する。代表例としては「特定の機関における文脈において、教師と学生に必要な、言語教育カリキュラムで必要な情報を収集し分析すること」(Brown, 2009, p. 269), 学習者のニーズに関する情報を収集する手

続き(Richards, 2001)などがある。ニーズ分析は元々はESP研究で実施してきた(Dudley-Evans & St John, 1998; West, 1994)。特定の職業やビジネスの現場などのジャンル別にどのような英語力が必要かを分析し、そのカリキュラムや教材の開発に繋げていくことを目的としていた。その後、Task-Based Language Teaching(TBLT, タスクを中心とした指導法)の分野において、どのような活動を実際に行うかタスクや教材を開発するために、ESPにおけるニーズ分析の知見が活用されてきた(Long, 2005)。

2.2.2 ニーズ分析の分類

ニーズ分析の構成要素には1. 調査対象・情報源(source), 2. 調査方法(method), 3. 分析単位(unit)の3種類が存在する(浦野, 2005, 2014)。1. 調査対象・情報源は誰からデータを得るか(教師・学生・卒業生・分野の専門家・教材)を表す。2. 調査方法はアンケートや観察、インタビュー、コーパスなどの言語使用に関するデータを収集する方法を指す。3. 分析単位は収集したデータを言語項目(語彙、文法)やスキル(例:リーディング、スピーキング), タスク(具体的な活動)のどの単位で分析するかを示す。ニーズ分析では「情報源と方法の組み合わせによるトライアングュレーション(source x method triangulation)」(Serafini, Lake, & Long, 2015), すなわち複数の情報源と調査方法を組み合わせたデータ収集・分析を行う必要がある。

調査対象・情報源は大きく分けて目標状況分析(Target Situation Analysis), 学習状況分析(Learning Situation Analysis), 現在状況分析(Present Situation Analysis)の3種類に分類される(Dudley-Evans & St John, 1998; 浦野, 2014)。目標状況分析は実際の使用場面で具体的にどのような活動を行うかという客観的なニーズを調査し、一方学習状況分析は学習者が学びたいことや不足していると感じていることなどの主観的なニーズを聞く。現在状況分析は目標状況分析と学習状況分析を比較した結果、現在何が足りていないかというギャップを明らかにする。

分析単位に関しては、ニーズ分析の結果をカリキュラム開発・教材開発等に繋げていくためには、言語項目(語彙、文法)やスキル(例:リード

ティング、スピーキング)ではなくタスクを分析単位とするのが望ましいとされる(浦野, 2005; 2014)。浦野(2005)はその理由はタスクを単位とした方法ならば言語の専門家でなくとも調査参加者が回答しやすいからだと述べている。また、浦野(2014)は日本における過去のニーズ分析の先行研究をレビューした結果、調査対象・情報源は学習状況分析が多く、分析単位がタスクのものが少なかったことを受け、ニーズ分析の結果を教材・カリキュラム開発など次に繋げるためには目標状況分析とタスクの組み合わせが必要だと主張している。

2.2.3 タスク型のニーズ分析

2.2.3.1 タスクの定義

タスクの定義は複数存在する。TBLTの文脈で使われる例は「毎日の生活で人が働いたり遊んだりする際に行うことすべて」(Long, 1985, p. 89)を指す。TBLTよりも広い定義としてはBachman and Palmer(1996)の「特定の状況で、何かしらのゴールや結果を達成するために、言語使用をともなう活動」(p. 447)がある。これら2つの定義をまとめると、タスクとは言語を用いて行う、目的を持った具体的な行動であると考えられる。

EMIにおいても例えば文法や英語表現など英語そのものを学ぶのではなく、実際に英語を用いて何かをすることから、EMIにおける活動もタスクとみなせると見える。Ortega(2015)はCLILとTBLTは開発してきた経緯が違うものの活動が似ていると述べている。EMIとCLILが似ていることからEMIとTBLTの相性がよく、そのためEMIのニーズ分析においてもTBLTで行われているようなタスクを分析単位とするのが有益だと考えられる。

以上を踏まえて、本研究におけるEMIで経験するタスクは「大学における英語による専門科目を学ぶにあたって、目的を持って何かを成し遂げるための活動すべて」と定義した。なお、ここでいうタスクは授業中に教師が学生に課す教室内活動ではなく、教室内外において学生が直面するすべての英語に関する目的を持った活動を指す。このためタスクは学生同士のインタラクティブなものだけではなく、1人で行うものも含まれる。

2.2.3.2 タスクを分析単位としたニーズ分析の先行研究

以下、EMIにも適用できると考えられるタスク型のニーズ分析のレビューを行う。収集したタスクの情報を元にアンケートを用いることで定量的なデータを集めることができ、カリキュラム・教材開発などの意思決定に使用することができる。具体的な方法としては、インタビューや観察により収集した具体的なタスクの記述(task statement)を作成し、このタスク一覧を元にリッカート法のアンケートを作成し、実際にタスクを経験する人に各タスクの頻度(frequency), 難度(difficulty), 重要度(importance)などを4-6段階で回答してもらう。その結果を用いて、各タスクがどの程度頻繁に遭遇するか・難しいか・重要かという情報を集めて、それらを意思決定の判断材料にする。この情報を元に、例えば頻度の高い場面・用例を収集した教科書を作成するトップダウン型の教育や、TBLTで教室用に作成したタスクを学生に1つずつ経験してもらうなどのボトムアップ型の教材開発が可能である(浦野, 2014)。

タスクを分析単位としたアンケートの先行研究としてはSerafini and Torres(2015), Lambert(2010), Rosenfeld, Leung, and Oltman(2001)がある。Serafini and Torresはスペイン語のビジネス場面で経験するタスク一覧やアンケートを作成し、アンケートの各タスクの頻度・難度の数値から高頻度・高難度・低頻度・低難度のタスクを分類した。分類の結果から高頻度・高難度のタスクなどの優先的に強化するタスクをまとめあげ、最終的にビジネスで5つの代表的な目標タスク(Target Task Types)に分類し、将来のTBLTのカリキュラム開発に活用した。Rosenfeld et al.はこれらを当時の最新のTOEFLテストに必要なタスクの選定に活用した。Lambertは特定の英語使用の目的がない日本の大学の英語カリキュラム内でTBLTを実施する際にどのようなタスクがふさわしいかを把握した。これらのよう、タスクを分析単位とした客観的なニーズを中心とした分析結果を、カリキュラム開発等の意思決定に活用することが可能である。

リッカート法におけるアンケートを作成する際に、各タスクの頻度・難度・重要度のどの情報

を回答してもらうかはニーズ分析の目的によって決まる。日本の大学におけるEMIのニーズ分析においてはまだ十分に情報がなく、学生がどのようなタスクを経験するか、またどのようなタスクに困難を覚えるかという情報が不足しているため、タスクの頻度・難度の情報が必要だと考えられる。

2.2.4 ニーズ分析の手法と

トライアングュレーション, 混合研究法

Serafini, Lake, and Long(2015)はニーズ分析の重要な貢献として方法論的示唆を提示することだと述べる。これが意味するところは以下の通りである。ニーズ分析の研究結果は厳密にはその結果を行った特定の場所におけるカリキュラム・教材開発等にしか使用できない。しかし、そこで得られたデータそのものは結果を行った箇所でしか使えなくとも、どのような方法論でニーズ分析を行ったかという情報は他者にも活用できるものである。このことから、どのような方法論で日本のEMIにおけるニーズ分析を行ったかを公表することは、将来の他のEMI環境におけるニーズ分析を行う他者にも有用であると考えられる。

本研究は特定の学校や機関のカリキュラム・教材開発を目指したものではない。それでも、ニーズ分析によって「学生の学習結果(Student Learning Outcome)」(Brown, 2016, p. 71), つまりコースを修了するまでにどのようなスキルや能力を身につけている必要があるかという最終成果に対する一般的な情報を得ることは可能である。この情報を元に、「日本のEMIにとって必要な英語力は何か」という疑問に対しても、タスクを分析単位としたニーズ分析結果によって、EMIでは「『〇〇が出来る』ことが必要である」と述べることが可能になる。より具体的には、EMIを専門科目の学習のみのコースとみなす立場ならEMI開始前までの準備コース等における必要な英語力の基準が示せる。EMIが言語学習も目的に含む立場ならEMIコースの過程における英語力の到達目標を示すことになる。いずれの立場にせよ、これらの情報がEMIを受講するにあたってどういった英語力を身につければいいかの判断に役立つ。

ニーズ分析は複数の情報源と方法からデータ収集・分析することが重要であるため(浦野, 2005; Long, 2005), 必然的に量的データ・質的データを組み合わせる混合研究法(Mixed-Methods Research)が必要となってくる。ニーズ分析における混合研究法はデータの解釈の質を向上させるのに役立つ(Brown, 2016)。同時に, Serafini, Lake, and Long(2015)は過去の包括的なレビューから推奨するニーズ分析の手順について述べている。情報源は実際にタスクを経験する内部の人間(例:現役学生)と外部の人間(例:卒業生・修了生)を両方含むこと, 方法はインタビューとアンケートという複数の手法を組み合わせることを推奨している。この方法を組み合わせる際に混合研究法が必要になってくることが多い。本研究の日本の大学におけるEMIにおけるニーズ分析では, 特に学習者からの視点という意味ではあまり先行研究が少ないので探索的な調査が必要であることと, 同時に量的なデータによって客観的なニーズを知ることが必要であるため, 質的・量的の両方のデータを用いて統合する必要がある。

2.2.5 本研究における研究質問 (Research Questions)

以上の先行研究を踏まえて, 本研究では以下の研究質問(Research Questions)を設定する。

日本の大学におけるEMIにおいて,

研究質問1: 学生はどのようなタスクを経験するか。

研究質問2: EMIの実施形態におけるタスクの頻度・難度の共通点と相違点は何か。

研究質問3: タスクの頻度・難度の情報と, 学生のEMIに対する受け止め方を統合して判明することは何か。

3 研究方法

3.1 研究方法概要

3.1.1 ニーズ分析の推奨される実施手順

本研究では日本の大学におけるEMIに必要な英語力を調査するためにニーズ分析を実施す

■表2: ニーズ分析において推奨される手順と本研究の研究手順

手順	推奨される実施方法	対応する本研究の手法
STEP 0	教育機関がコミュニケーションにおける問題を特定	なし
STEP 1	実際に学んでいる学習者を対象に半構造化インタビューを実施	インタビュー調査
STEP 1A	インタビューの予備調査の質問を作り目標タスクを決定, 必要に応じてフィードバックを受けて修正	予備調査を実施
STEP 1B	インタビューを実施し分析	1対1インタビュー, フォーカスグループインタビュー
STEP 2	インタビューの結果がより広い目標の母集団にも当てはまるか一般性を確認	アンケート調査
STEP 2A	インタビューの結果を元にアンケートの項目を作成	GTxAによるアンケート項目作成
STEP 2B	予備調査のアンケートを作成, 必要に応じて修正	予備調査を実施
STEP 2C	アンケートを実際の集団を相手に実施	アンケート実施
STEP 3	追加の道具や手続きを用いてフォローアップデータを収集	アンケートの自由回答記述データ分析
STEP 4	問題のある目標タスクを特定するために, 複数のデータ源と方法による結果のトライアングュレーションを行う	量的・質的データの統合・解釈

(注1) Serafini, Lake, and Long(2015)を元に作成。(注2) GTxA = Grounded Text Mining Approach(詳細は3.2参照)

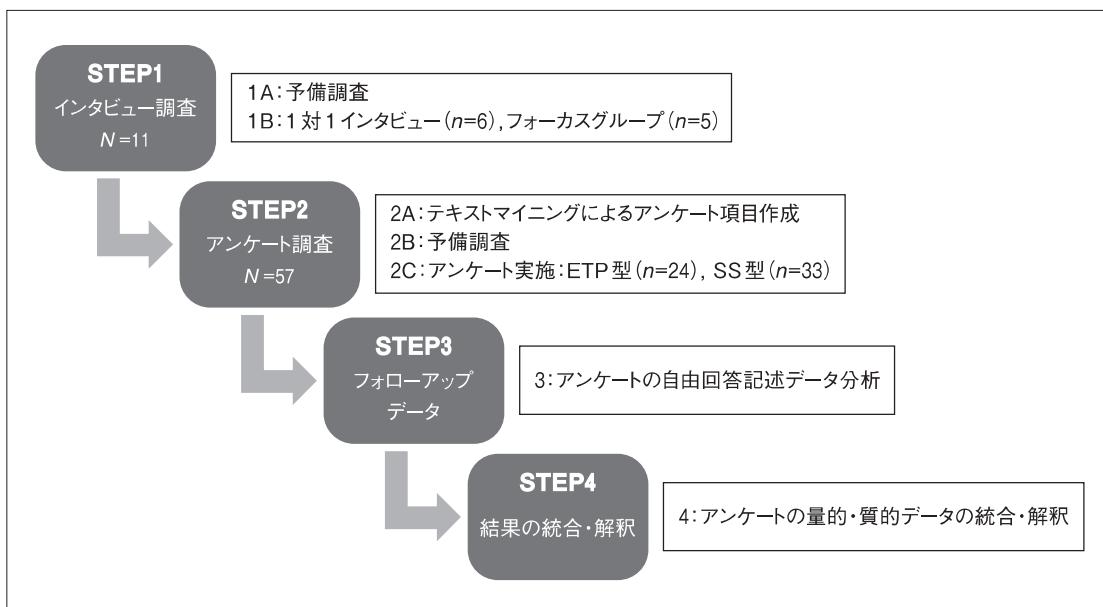
る。本研究は Serafini, Lake, and Long (2015) が提唱するニーズ分析の手順に従った。表2は Serafini et al. が推奨する手順と、本研究における手順を対応させたものを表す。

今回は複数の段階を含む。最初にインタビューデータを収集して分析を行い、その結果を元にアンケートを作成する。次にアンケートの実施を行い、それらの結果とインタビューデータ・自

由回答記述データとの比較・対照を行い、最終的な結論を作成する。

3.1.2 本研究の手続き的ダイアグラム

Serafini, Lake, & Long (2015) が推奨するニーズ分析の手法を元に、本研究のデータ収集・分析を視覚的に表す手続き的ダイアグラム作成した(図1)。



■図1:本研究における手続き的ダイアグラム

(注)図中のSTEPは表2中のニーズ分析の推奨される手順に対応している。

各STEPの詳細は以下の通りである。

- STEP1: 予備調査の後、参加者に1対1インタビュー(6名)およびフォーカスグループインタビュー(5名1組)を実施する。
- STEP2: 文字化されたインタビューデータをテキストマイニングの結果を元に分析し、各タスクの頻度・難易度に関する情報をアンケート作成する。予備調査の後、作成したアンケートをEMI受講経験のある現役学生と卒業生・修了生計57名に実施する。アンケートから得られた各タスクの頻度(高・低)と難易度(高・低)の平均値を元に4分類しマッピングする。
- STEP3: 自由回答記述データの分析を行い、各タスクを頻度・難度を4種類に分類し

た結果を補足・説明する参加者の回答を抽出・検討する。

- STEP4: アンケートおよび自由回答記述から得られた各データを統合・解釈し、研究結果全体から言える結論を導き出す。

3.2 GTxA (Grounded Text Mining Approach)

本研究は混合研究法を採用している。混合研究法は量的データと質的データのどちらかだけでは見えてこないものを組み合わせることによって、どちらか一方の時よりも研究課題に関するより良い理解を目指すアプローチを指す(Fetters & Freshwater, 2015)。前述の通り、ニーズ分析の把握には「情報源と方法の組み合わせによるトライアンギュレーション」(Serafini, Lake, &

Long, 2015) が大事であり、そのために複数のデータから把握する混合研究法を実施した。

また、本研究は混合研究法のアプローチとして Grounded Text Mining Approach (以下 GTxA, Inaba & Kakai, 2019) を採用した。GTxA はテキストデータを分析する際に研究者による主観的な判断とテキストマイニングによる客観的な分析を組み合わせる手法である。具体的には、GTxA はまずテキストマイニングの出力結果によって質的データの感じを数量的・視覚的に把握し、その情報を元に研究者が手動で質的分析を実施する。Inaba and Kakai はテキストマイニングと手動の分析の組み合わせによって片方の分析では得られないデータの理解が得られ、同時に大量のデータを分析する際の認知的負荷の軽減が可能だと述べる。

本研究が GTxA を採用する理由は、この手法がニーズ分析を効率化することにある。インタビューデータからアンケート項目を作成する際に、テキストマイニングの出力データはタスクのキーワードを特定し、そのキーワードを元にインタビューの文字起こしデータの該当箇所を読む、探すことなどの手続きを容易にする。アンケートにおいて収集した自由回答記述データを

分析する際にもテキストマイニングで得た量的データを元に質的データの分析を効率的に行うことができる。このように、膨大なデータからニーズ分析に必要な情報を短期間に効果的に取得するのに GTxA は大きな役割を果たす。

3.3 研究協力者と場所

3.3以降では各 STEP における研究協力者や場所、データ収集・データ分析の手順の詳細を述べる。本研究は日本の私立大学における5箇所の学部・大学院を対象とした(表3)。このプログラムで EMI を経験した現役学生と卒業生・修了生を対象にデータ収集を実施した。そのうち4つは学部・大学院が対応している2箇所ともう一つは独立した学部のプログラムであった。A と B, D と E はそれぞれ同じ箇所にあり接続されている学部・大学院である。この5箇所の学部・大学院は Brown and Iyobe(2014) の EMI の6分類に照らし合わせると、主に “2. Semi-Structured” 型(以下 SS型) と “5. English-Taught Program” 型(以下 ETP型) の2種類に分かれた。インタビューは C, D, E のプログラム、アンケートは全てのプログラムの学生と卒業生・修了生が参加した。

■表3: データ収集を行った学部・大学院の情報

プログラム名	学部 / 大学院	専攻	EMI 分類	インタビュー	アンケート回答
A	学部	リベラルアーツ	ETP型	不参加	参加
B	大学院	リベラルアーツ	ETP型	不参加	参加
C	学部	国際日本文化	ETP型	参加	参加
D	学部	英語教育・英語学・英米文学	SS型	参加	参加
E	大学院	英語教育・英語学・英米文学	SS型	参加	参加

3.3.1 インタビュー参加者・アンケート回答者

インタビューは表3における ETP型が3人(プログラム C), SS型が8名(プログラム D, E)の合計11名が参加した。アンケートの回答者は ETP型 が24名(プログラム A, B, C), SS型が33名(プログラム D, E) が回答した。

3.3.2 研究倫理

調査を実施する私立大学の研究倫理審査委員会から承認を得てデータ収集を実施した。インタビュー・アンケート両方の参加者に、データの匿名化・特定の授業の単位取得や学業成績に関わることはないことなどの旨を伝え、研究参加者からの了承を得てデータ収集を行った。

3.4 データ収集

3.4.1 インタビューデータ

インタビューは1対1インタビューとフォーカスグループインタビューの2種類を実施した。1対1インタビューは各研究者1名と参加者1名で実施した。フォーカスグループインタビューは特定のテーマに関して複数人のグループで落ち着いた雰囲気でディスカッションを行う形式の調査方法である(Krueger & Casey, 2009)。フォーカスグループインタビューの特徴は相乗効果性・雪だるま性・刺激性・安心感・自発性の5つである(Hess, 1968, in ヴォーン, シューム, & シナグブ, 1996=1999, p. 20)。つまり、周りの発話に反応してどんどん会話が広がっていく(相乗効果性・雪だるま性・刺激性)のと、必ずしもすべての質問に答えることなく、自分の話したいタイミングで話せること(安心性・自発性)が特徴である。このような情報は1対1インタビューだけでは得られない種類のデータとなる。ニーズ分析は方法のトライアングュレーションが必要であるため1対1インタビューに加えて実施した。

1対1インタビューは2020年1月から2月の間に実施し、参加者は6名であった。1対1インタビューは共同研究者を含む3名で分担して実施した。6名のインタビュー時間の合計は474分であった。フォーカスグループインタビューは2020年2月に実施した。フォーカスグループインタビューは共同研究者3名と研究補助者1名の計4名で、参加者5名と共に実施した。フォーカスグループのインタビュー時間は124分であった。

1対1・フォーカスグループのいずれのインタビューにおいても、参加者は事前にGoogleフォーム上で年齢・学年・所属・入学年度・受講したEMI科目について記入した。1対1インタビューではそれらの情報を元にEMIで経験したタスクの種類や頻度・難度について答えてもらった。必要に応じてEMIに関してどのように考えているかについても述べてもらった。1対1インタビューはある大学の教室において実施を行った。1対1インタビューはレコーダーによって録音を行った。

フォーカスグループインタビューではEMIで経験したタスクの種類や頻度・難度について参加

者5名で自由に話しあってもらった。必要に応じてその場で司会を行っていた研究代表者が質問を行った。フォーカスグループインタビューはビデオカメラとレコーダーを用いて録画・録音を行った。録画データは発話者の特定や意図の確認、フォーカスグループの雰囲気の記録のためを使用した。

3.4.2 アンケートデータ(量的)の収集

表3の学部・大学院に対応するEMIを経験した現役学生と卒業生・修了生にGoogleフォームで回答を依頼した。データ収集は2020年5月に実施した。回答者は最初に名前、年齢、学年、入学年度、所属(学部・学科、卒業生・修了生は最後に所属した学部・研究科)、海外滞在経験、英語力の資格試験のスコアを基礎情報として記入した。その後に、インタビューデータから作成したEMIで経験した各タスクに関してリッカート式のアンケートを回答してもらった(3.5.2.1参照)。EMI授業を受けてきた全体的な印象を元に、経験してきた各タスクに関して頻度・難度を「高い」「やや高い」「やや低い」「低い」の4件法で回答してもらった。また、経験がないタスクに関しては「経験なし」を回答してもらうように依頼した。

3.4.3 自由回答記述データ(質的)の収集

上のリッカート式のアンケートに加えて自由回答記述で内容を入力してもらった。質問内容は以下の3点である。

- 「質問1. 英語による専門科目(EMI)の授業において、今よりも効率良く学ぶためにはどういった能力が必要だと思いますか。その理由も合わせて教えて下さい。卒業生・修了生の方は在学時にどういった能力がもっと必要だったかを教えて下さい。」
- 「質問2. 英語による専門科目(EMI)の授業を受講するにあたって、どのような授業やサポートがあればより受講しやすくなるか、その理由も合わせて教えて下さい。卒業生・修了生の方は在学時にどのような授業やサポートがあれば良かったかを教えて下さい。」
- 「質問3. 最後に、EMIに関して何かコメントがありましたらお書き下さい。」

3.5 データ分析

3.5.1 インタビューデータの分析

1対1インタビューデータの文字起こしは一部は共同研究者で分担して行い、残りは文字起こし業者に委託を行った。フォーカスグループインタビューは研究代表者が録画したデータとフィールドノーツ、ホワイトボードのメモを確認しながら文字起こしと分析を実施した。

本研究のGTXAにおけるテキストマイニングの際にはKH Coder(樋口, 2004)を使用した。KH Coderはテキストファイルを読み込むとデータからキーワードを自動で抽出するフリーソフトウェアである。KH Coderは目的に応じてその結果を複数の形で出力できる。本研究では抽出語リストや共起ネットワーク図、対応分析図等を出力し、3.5.2.1のタスク表作成・アンケート作成や3.5.3の自由回答記述データの分析に活用した(実際の出力された表・図の例は巻末の資料1と資料2、資料7を参照)。

3.5.2 アンケートデータの分析(量的)

3.5.2.1 タスク一覧とアンケートの作成・実施

テキストマイニングによって出力された抽出語リストと共にネットワーク図(巻末資料1と資料2参照)からEMIにおけるタスクに関するキーワードを抽出し、それと対応する語が含まれる一文をインタビューデータから探し出し、それを元にタスクの記述(task statement)を1つ1つ作成した。共同研究者3名で会議を行いながらタスクを作成する作業を複数回繰り返し、タスク一覧を作成した。5名の研究協力者によるフィードバックを受けた後、最終的なタスク一覧を完成させた。このタスク一覧を元に、各タスクの頻度・難度について回答するアンケートを作成した。「高い」「やや高い」「やや低い」「低い」の回答をそれぞれ“4”, “3”, “2”, “1”と数値に変換し、「経験なし」は“0”とみなして計算を行った。

3.5.2.2 各タスクの頻度・難度による分類

これらを元にアンケートを実施した結果、頻度・難度の高低によってタスクの分類を行った。ETP型とSS型のそれぞれのデータにおいて、頻度の高・低と難度の高・低の組み合わせの合計4種

類に分類した。頻度・難度の高低の基準はそれぞれの全体の平均値を基準とした。各タスクの平均値を頻度を縦軸、難度を横軸とした座標で表し、それぞれ高頻度・高難度(図中右上)、高頻度・低難度(図中左上)、低頻度・低難度(図中左下)、低頻度・高難度(図中右下)の4種類に分類した(実際の図は図2・図3参照)。

3.5.3 自由回答記述データの分析(質的)

自由回答記述データに対してテーマ分析を行った。テーマ分析はデータからパターンを見つける、分析し、報告する質的研究の分析手法である(Braun & Clarke, 2006)。本研究はGoogleフォームの自由回答記述欄で参加者に「EMIに必要な能力」「EMIに必要なサポート」に関して聞いているため、この2つのカテゴリーごとに複数回現れたパターンを探しだした。このパターンを探す際にもGTXAを使用し、テキストマイニングの結果から2つのカテゴリーにおけるキーワードを特定し、それを元にサブカテゴリーとなるパターンを探す作業を行った。

4 結果

本研究はインタビューデータを元にEMIで経験するタスク一覧を明らかにし、それを元に作成した各タスクの頻度・難度を4段階で尋ねるリットカート式のアンケートを実施した。アンケートでは同時に自由回答記述形式でEMIに必要な能力やサポートに対してどのように考えていたかを調査した。

4.1 タスク表の作成と アンケートデータの分析結果

4.1.1 タスク一覧

テキストマイニングの出力結果の抽出語リストと対応ネットワーク図(巻末資料2参照)を元にタスクを表すキーワードを決め、そのキーワードを元にインタビューの文字起こしデータからタスクの記述(task statement)を1つ1つ作成した。インタビューデータを包括的に検索し終わるまでタスク作成を続け、最終的にEMIで経験

するタスク一覧を作成した(表4)。タスクは合計で47個特定され、これを「授業前後」「レクチャー」の5種類のカテゴリーに分類した。

■表4: EMIで経験したタスク一覧

カテゴリー	タスク一覧
1. 授業前後	1. 専門科目の内容をインターネットで英語で調べる 2. 英語で書かれた専門の教科書を読む際に分からない単語を調べる 3. 英語で書かれた専門の教科書の指定された章を読む 4. 英語で書かれた専門の教科書の概要・大意を把握する 5. 英語で書かれた専門の教科書の中から特定の情報を探し出す 6. 専門内容を理解するために英語で書かれた専門の教科書をじっくり読む 7. 複数の英語で書かれた文献を読んで内容を比較する 8. 英語で書かれたリーディング課題を和訳する 9. 専門科目の内容について質問するために先生に英語でメールを送る 10. 専門授業に関して交渉するために先生に英語でメールを送る
	11. 専門用語を英語で定義し例を挙げて書く 12. 先生から専門内容についての説明を英語で聞く 13. 英語で書かれた専門の教科書の重要な点を英語で書く 14. 先生からの専門内容についての説明をノートに英語で書く 15. 先生に専門内容について英語で質問する
	16. 先生の意見に対して先生に英語で反論する 17. 専門内容に関して英語教材(音声・映像)を視聴する 18. 授業で使用される、英語で書かれた資料(ハンドアウト・スライド)の内容を理解する 19. 他の学生が書いた英語の作品を英語で口頭でフィードバックする 20. 日本語で書かれた資料について英語で口頭で質問する 21. 教場テストでエッセイを英語で書く
	22. 英語で書かれた専門の教科書で読んだ内容についてディスカッションをする 23. 与えられた質問に対しペアで英語でディスカッションする 24. ディスカッション中に専門内容について英語で説明する 25. ディスカッション中に専門内容について他の学生に英語で質問をする 26. ディスカッション中に他の学生に発言を英語で促す 27. グループ代表としてグループディスカッションの内容をクラス全体に紹介する 28. 専門内容について英語でクラス全体でディスカッションをする
	29. 専門科目の内容についてプレゼンテーションのスライドを英語で書く 30. 英語で書かれた専門の教科書で読んだ内容のまとめをプレゼンテーションする 31. 専門科目の内容についてグループで英語でプレゼンテーションをする 32. プrezentationでディスカッションクエスチョンを英語で投げかける 33. 自分の発表した内容について先生や学生からの意見を英語で聞く 34. 他の学生から専門内容についての説明を英語で聞く 35. プrezenterが投げかけたディスカッションクエスチョンに関してグループで話し合う 36. プrezentationの内容について口頭でフィードバックをする 37. 学生のプレゼンテーションについて評価シート等でフィードバックを英語で書く
	38. 日本語で書かれた資料について英語で書く 39. 英語で書かれた専門の教科書の要約を英語で書く 40. 英語で書かれた専門の教科書で読んだ内容について自分の考察を英語で書く 41. レポートで引用するために参考文献を英語で読む 42. 専門科目の内容についてレポートを英語で書く 43. 専門科目の内容についてグループで分担して英語でレポートを書く 44. 英語のライティング課題を他の学生からフィードバックを受けて修正して提出する 45. 専門内容の知識を利用して英語で創作をする(例:小説・詩・演劇, etc...) 46. 専門科目の内容について学期末レポートを英語で書く 47. 学位論文(卒業論文や修士論文)を英語で書く
2. レクチャー	
3. ディスカッション	
4. プrezentation	
5. 課題	

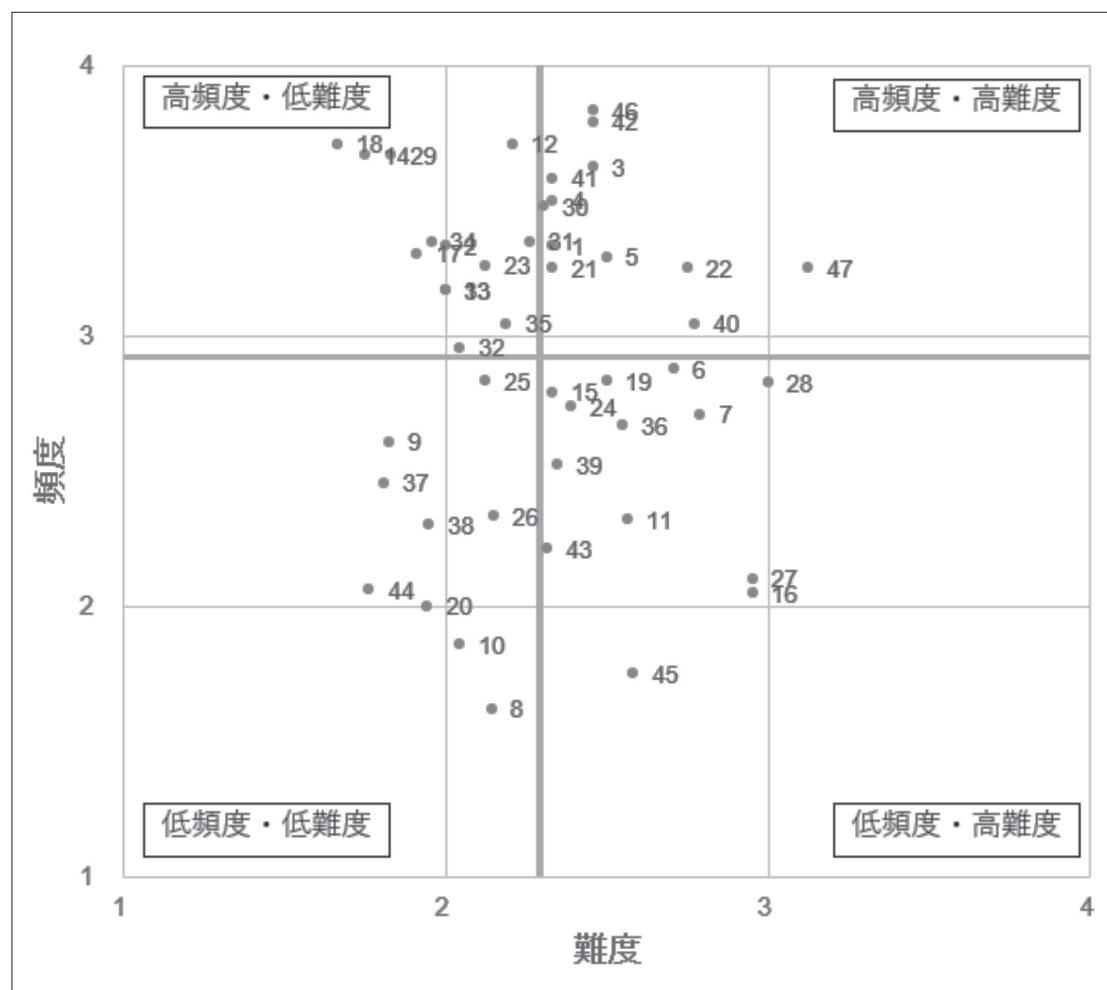
4.1.2 アンケート結果(量的)

このタスク一覧を元にタスクの頻度・難度を4件法で聞くリッカート式のアンケートを作成し、ETP型とSS型のEMIを経験した学生・卒業生(修了生)に回答してもらった。ETP型、SS型の各タスクにおいて、頻度・難度を「高い」「やや高い」「やや低い」「低い」「経験なし」のどれを選択したかを数値化するために、頻度・難度共に「4=高い」、「3=やや高い」、「2=やや低い」、「1=低い」に変換した(平均値は「経験なし」を0とみなした)。これらの数値を用いて、各タスクごとに頻度・難度の平均値を計算した。(具体的な頻度・難度のパーセンテージと平均値は巻末の資料3・資料4を参照)。

4.2 各タスクの頻度・難度別の4分類

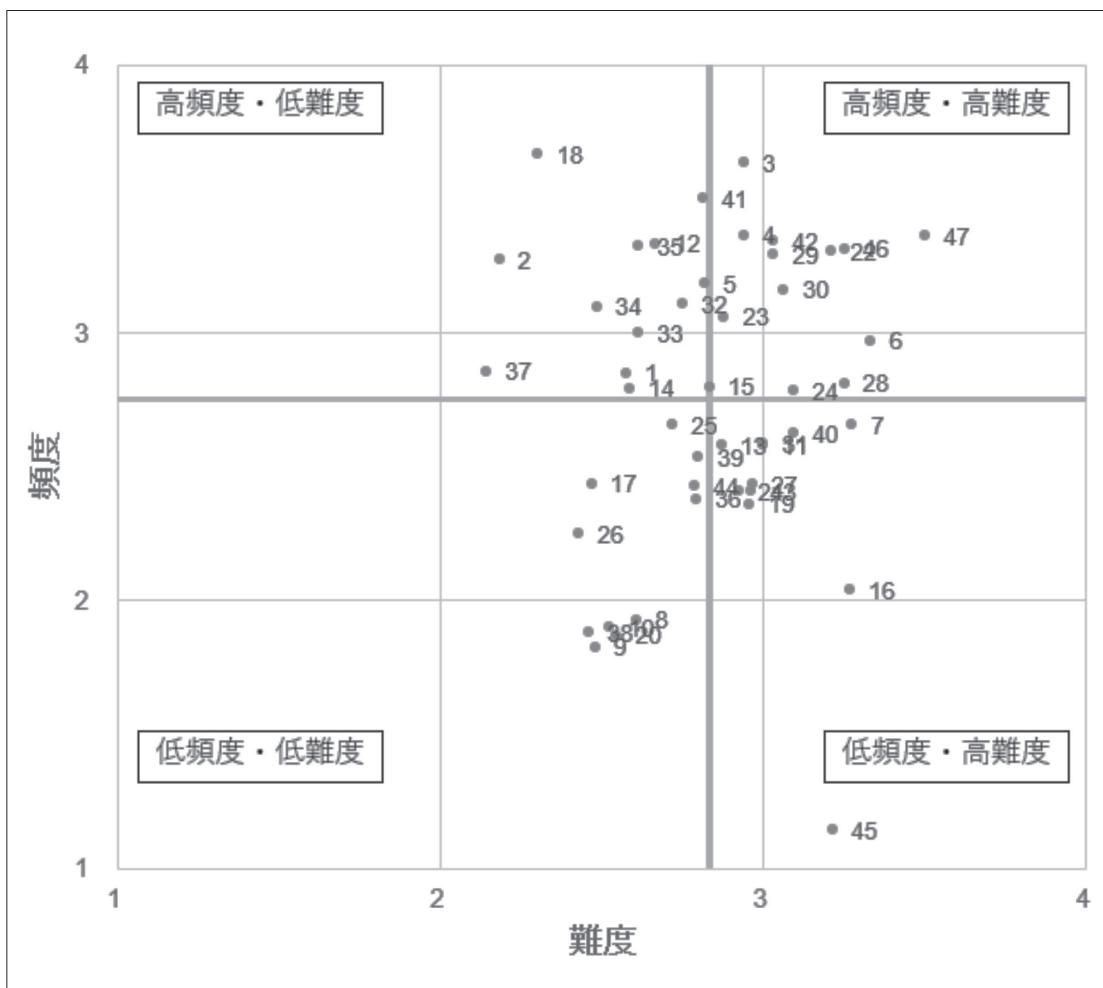
4.2.1 各タスクの分類を表す4分類のマッピング

4.1で明らかになった各タスクの頻度・難度の平均値を元に分類を行った。ETP型・SS型のそれぞれのEMIの実施形態ごとに、頻度(高・低)と難度(高・低)組み合わせの合計4種類に分類した。図2と図3はETP型、SS型別の各タスクの頻度・難度の平均値をマッピングしたものである。頻度・難度の平均値を基準として高低として、高頻度・高難度(図中右上)、高頻度・低難度(図中左上)、低頻度・低難度(図中左下)、低頻度・高難度(図中右下)の4種類に分類した。



■図2:タスク別頻度・難度4分類マッピング(ETP型)

(注)縦軸は頻度(1-4)、横軸は難度(1-4)の値を表す。各ドットの番号は表4のタスク番号を表す。
図中の太線は頻度と難度の全体の平均値を表す(頻度平均値=2.92、難度平均値=2.29)。



■図3:タスク別頻度・難度4分類マッピング (SS型)

(注)縦軸は頻度(1-4), 横軸は難度(1-4)の値を表す。各ドットの番号は表4のタスク番号を表す。

図中の太線は頻度と難度の全体の平均値を表す(頻度平均値=2.75, 難度平均値=2.83)。

4.2.2 ETP型とSS型における各タスクの カテゴリー別の4分類の比較・対照

4.2.1で座標で表した4分類の各タスクの分布の位置がわかるように、表5でETP型, SS型における各カテゴリーのタスクの分類を4種類別で行った。

4.2.2.1 ETP型・SS型の共通点

ETP型・SS型の両EMI実施形態で共に高頻度・高難度だったのはタスク番号3, 4, 22, 30, 46, 47だった(タスク番号は表4を参照)。これらは大きく3種類に分かれる。「3. 英語で書かれた専門の教科書の指定された章を読む」「4. 英語で書かれた専門の教科書の概要・大意を把握する」は教科書のリーディング能力を表す。「22. 英語で書

かれた専門の教科書で読んだ内容についてディスカッションをする」「30. 英語で書かれた専門の教科書で読んだ内容のまとめをプレゼンテーションする」は読んだ内容を書いたり発言する際に統合する技能統合(integration)の能力を表す。「46. 専門科目の内容について学期末レポートを英語で書く」「47. 学位論文(卒業論文や修士論文)を英語で書く」は長い論文を書く能力を表す。

一方、ETP型・SS型で共に低頻度・低難度はタスク番号8, 9, 10, 20, 25, 26, 38, 44であった。44を除く7つのタスクは以下の3種類に分類できる。「9. 専門科目の内容について質問するために先生に英語でメールを送る」「10. 専門授業に関して交渉するために先生に英語でメールを送る」はメールを

■表5: ETP型・SS型におけるカテゴリー別の4分類の各タスクの分布

	高頻度・高難度 (図中右上)	高頻度・低難度 (図中左上)	低頻度・低難度 (図中左下)	低頻度・高難度 (図中右下)
ETP型				
1. 授業前後	1, 3, 4, 5	2	8, 9, 10	6, 7
2. レクチャー	21	12, 13, 14, 17, 18	20	11, 15, 16, 19
3. ディスカッション	22	23	25, 26	24, 27, 28
4. プレゼンテーション	30	29, 31, 32, 33, 34, 35	37	36
5. 課題	40, 41, 42, 46, 47		38, 44	39, 43, 45
SS型				
1. 授業前後	3, 4, 6	1, 2, 5	8, 9, 10	7
2. レクチャー	15	12, 14, 18	17, 20	11, 13, 16, 19, 21
3. ディスカッション	22, 23, 24, 28		25, 26	27
4. プレゼンテーション	29, 30	32, 33, 34, 35, 37	36	31
5. 課題	42, 46, 47	41	38, 39, 44	40, 43, 45

(注)「図中」の「図」は図2と図3のマッピングを表す。

送付する行為を指す。「8. 英語で書かれたリーディング課題を和訳する」「20. 日本語で書かれた資料について英語で口頭で質問する」「38. 日本語で書かれた資料について英語で書く」は日本語資料の使用を指す。「25. ディスカッション中に専門内容について他の学生に英語で質問をする」「26. ディスカッション中に他の学生に発言を英語で促す」は教室内で友人に質問を投げかける行為を指す。

4.2.2.2 ETP型・SS型の相違点

ETP型とSS型の相違点は、まず、表5のカテゴリー別の分類を見ると、ディスカッションに関するタスクはSS型が高頻度・高難度のタスクが4種類あるのに対し、ETPは22のみである。ETP型では高頻度・低難度だったのは「23. 与えられた質問に対しペアで英語でディスカッションする」、低頻度・高難度であったのは「24. ディスカッション中に専門内容について英語で説明する」「28. 専門内容について英語でクラス全体でディスカッションをする」である。

また、全体的なマッピングの分布を見ると、SS型の学生のほうが、ETP型よりも全体的にタスクの難易度が高かった。

4.2.3 自由回答記述データの分析

自由回答記述データのテーマ分析を実施した。

自由回答記述データのテキストマイニングで分析した結果、「必要な能力」「必要なサポート」の2種類が中心になっていることを確認できた。「必要な能力」「必要なサポート」のそれぞれの回答で繰り返し現れるパターンを見つけ出し、サブカテゴリーを作成した。カテゴリーを探す際には、テキストマイニングの出力結果(巻末資料7参照)からキーワードを探し出し、そのキーワードを元に自由回答記述データを検索し分析を行った(各サブカテゴリーの内訳における具体的な回答の抜粋は資料5と資料6を参照)。

表6の通り、EMIに必要な能力の受け止め方は4つのサブカテゴリーに分類され、合計58回の言及があった。各サブカテゴリーの説明は以下の通りである。「技能・能力」は直接特定のスキルに言及していたものである。「知識・理解」は専門内容のインプットに関するものを指す。「授業内の発言」はアウトプット全般(スピーキング・ライティングのスキルや自分の意見を持つ・述べる等)に関するものであった。「情意面」はモチベーションや自信など情意的要因に関して述べたものであった。また表7の通り、EMIに必要なサポートの受け止め方はアカデミックスキル、専門知識・内容、理解・確認、学習環境の4種類のサブカテゴリーに分類され、合計64回の言及があった。

■表6: EMIに必要な能力に対するテーマ分析結果
(サブカテゴリーと内訳)

サブカテゴリー	内訳
技能・能力(30)	リーディング(12)
	スピーキング(7)
	リスニング(3)
	ライティング(2)
	その他(6)
知識・理解(11)	日本語による理解(6)
	専門知識(5)
授業内の発言(8)	アウトプット(4)
	発信力(4)
情意面(9)	情意面(9)

(注) サブカテゴリー、内訳内の数字はコメント数を表す(合計58個)。

■表7: EMIに必要なサポートに対するテーマ分析結果
(サブカテゴリーと内訳)

サブカテゴリー	内訳
アカデミックスキル(19)	全般(6)
	スピーキング(2)
	ライティング(6)
	使用機会(5)
専門知識・内容(12)	内容(4)
	日本語による補助(8)
理解・確認(9)	フィードバック(4)
	TAやピアによるサポート(5)
学習環境(24)	授業進行(13)
	雰囲気・設備(7)
	時間・精神的ゆとり(4)

(注) サブカテゴリー、内訳内の数字はコメント数を表す(合計64個)。

5 考察

本研究はインタビューデータから47個のEMIの代表的なタスクを発見し、それを元に各タスクの頻度・難度を4件法で尋ねるアンケートを作成・実施した。各タスクの頻度・難度の平均値を元にタスクを「高頻度・高難度」「高頻度・低難度」「低頻度・低難度」「低頻度・高難度」の4種類に分類した。その分類を元に、ETP型とSS型の各EMIの実施形態におけるタスクの分布に基づく共通点・相違点を明らかにした。同時に、自由回答記述データのテーマ分析によりEMIに対する必要な能力・必要なサポートの考え方を分析し、量的データの結果をより説明するデータを得られた。これらの結果の統合・解釈によって、EMIに必要な英語力を明らかにした。以下、2.2.5で述べた3つの研究質問に対する答えを述べていく。

5.1 研究質問1: 学生はどのようなタスクを経験するか

テキストマイニングの補助を元にインタビューデータを分析した結果、合計47個のタスクが抽出された(表4参照)。創作活動(タスク番号45)など「経験なし」のパーセンテージが高い(50%)こ

とから(巻末の資料3・資料4参照)一部コース特有のタスクも含まれている。しかし、これらは実際にEMIを経験した参加者からの声を元に、タスクの取りこぼしがないようにインタビューデータを分析した結果である。これらを踏まえると本研究の文脈におけるEMIの代表的なタスク一覧が完成したと言える。このタスク一覧は本研究のみならず、今後の日本におけるEMIの実証研究に活用できると考えられる。

5.2 研究質問2: EMIの実施形態におけるタスクの頻度・難度の共通点と相違点は何か

タスクの頻度・難度の4分類(表5参照)を元に、ETP型、SS型の共通点と相違点が明らかになった。

5.2.1 ETP型・SS型の共通点

ETP型・SS型の両EMI実施形態で共に高頻度・高難度のタスクはリーディング能力、技能統合能力、論文執筆能力の3種類であった。まとめると、「リーディング」「技能統合」「論文執筆」の3種類のタスクが出来ることが喫緊のEMIに必要な英語力だと言える。これはEMIは授業内言語が英語であり、オールイングリッシュでのリスニングやスピーキングが問題だという直感とは反対の結果である点が興味深いと言える。読む

能力が重要な点に関しては Sawaki(2017) の教員視点のニーズ分析の結果とも一致する。これらの情報を元に, Serafini and Torres(2015) がビジネス場面における典型的な目標タスクを作成したように、「文献を読み, それらを要約・引用しながらレポートを執筆する」などの代表的な場面を考えた教育用タスクが作成可能である。

一方, ETP型・SS型で共に低頻度・低難度のタスクはメールの送付, 日本語資料の使用, 友人に質問を投げかける行為の3種類であった。メールの送付に関してはSS型はそもそも「経験なし」が多かったため(資料2参照), 教員には日本語でメールを送付している可能性も考えられる。日本語使用に関してはEMIプログラムでは授業内は日本語をあまり活用しないことが伺える。一方, 「44. 英語のライティング課題を他の学生からフィードバックを受けて修正して提出する」はETP型のみ「経験なし」の回答が多い。これはETP型の学生がそもそもフィードバックを受ける機会が少ないと起因すると考えられる。これを裏づけるデータは5.3で考察を行う。

5.2.2 ETP型・SS型の相違点

ETP型とSS型の相違点は, まず表5のカテゴリー別の分類を見るとディスカッションのタスクがETP型とSS型で分布が異なる。この理由は両EMIの形態でディスカッション時の雰囲気が異なる可能性がある(自由回答記述データによる詳細は5.3参考)。また, 「6. 専門内容を理解するために英語で書かれた専門の教科書をじっくり読む」がETP型では低頻度・高難度で, SS型では高頻度・高難度である。この理由の1つには, ETP型とSS型の形態の差であり, ETP型はほぼすべての授業が英語で行われるため英語で読む量も多いのに対し, SS型は選択必修の授業が多く必ずしもすべての授業が英語ではないことがある。よって, SS型は速読に重きを置かない可能性がある(これを裏づけるデータは5.3参考)

5.3 研究質問3: タスクの頻度・難度の情報と, 学生のEMIに対する受け止め方を統合して判明することは何か

自由回答記述データをテーマ分析した結果, アンケートの量的データの結果をより良く説明

できる情報も得ることができた。まず, 前述の通りETP型は「44. 英語のライティング課題を他の学生からフィードバックを受けて修正して提出する」の「経験なし」の割合がSS型よりも多い。ETP型の自由回答記述は他者からのフィードバックに関するものが4点あった。そのうちの代表例が以下の通りである:

「日本の大学はレポートなどにフィードバックがないのが多く, 学生一人ひとりのきめ細やかなフィードバックがあれば, 自分の弱点も見えて勉強しやすくなるかなと留学して思った(ETP)」

「レポートを書いてもほとんどのクラス(特に大教室)で教授からのフィードバックがなかったため, 自分のレポートのどこが良かったのか, どこを改善すればいいのか全くわからなかった(ETP)」

この通り, 自分の提出したレポートのフィードバックそのものがなかったため, タスク番号44のフィードバックを受けて提出するタスクを経験する機会も少なかったと考えられる。ETP型・SS型の共通点で高頻度・高難度で長い論文を書くことが重要なことからも, 学生が自らのレポートの良い点・改善する点を知る機会を得ることがEMIプログラムにとって必要である。

また, 5.2.2で述べたとおり, ディスカッションのタスクがETP型とSS型で4分類の分布が異なる。その一因として, SS型の学生の複数の回答が得られた。

「質問する雰囲気が重たい, しづらい(SS)」

「教授と生徒が双方向で意見を言える環境があると, わからない部分を聞けたり自分も参加しているという意識ももてて理解力も上がると思う (SS)」

このような雰囲気になる理由として考えられるのは, SS型はEMIで行われる授業が選択必修科目であることが多く, 学生の中には普段は専門を日本語で学ぶことが中心であったり, そも

そもそも入学前にEMIを受講するプログラムであることを知らなかった可能性もある。SS型のEMI実施の際には、このディスカッションの雰囲気に対してサポートを行えることが重要である。

さらに、5.2.2で述べたとおり、ETP型ではリーディング課題をじっくり読む頻度がSS型よりも低い。それを示す回答としてETP型の学生による、EMIに必要な能力に関するコメントが複数あった。

「文献を素早く読み内容を取る能力(ETP)」
 「英語の論文を早く読む能力(ETP)」
 「文章処理のスピード(ETP)」
 「パラグラフごとの要約をつかみながら文章を読む能力。ある程度の速度で多くの分量を読む必要があったため。(ETP)」

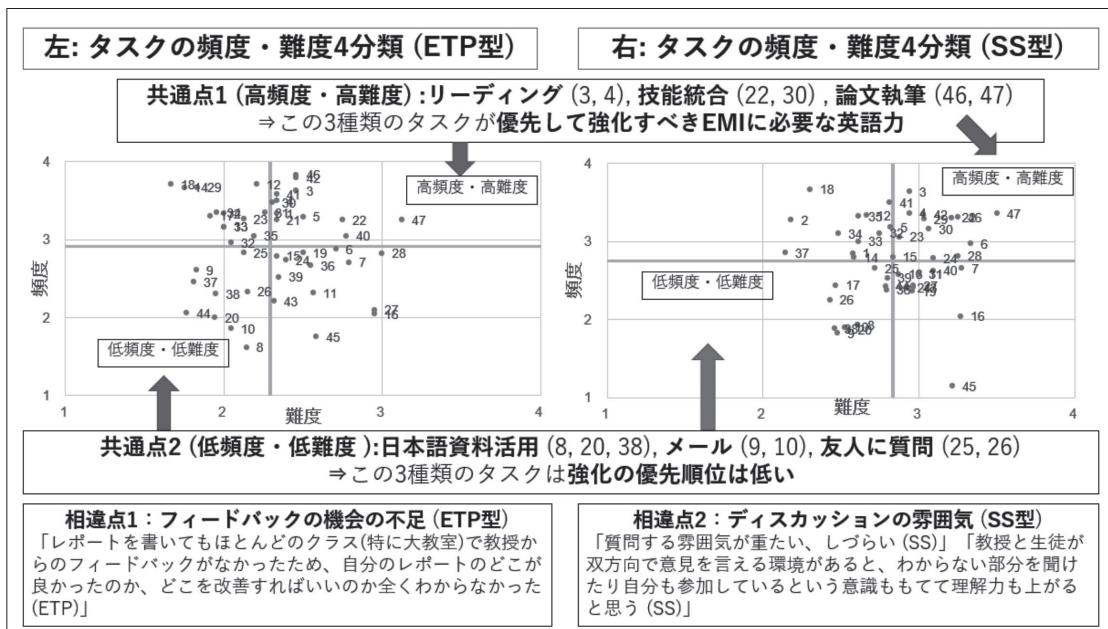
のことから、リーディング能力はETP型・SS型共に高頻度・高難度のタスクが2つあったが（タスク番号3・4）、ETP型のような英語の授業の割合が多い形態では特にこの素早くリーディングを行い、内容を把握する訓練が必要である。これらは高校までの英語学習で求められていたリーディング能力とは異なる可能性もあるため、EMIを受

講するにあたり訓練が必要な点であると言える。

6 結論

6.1 本研究のまとめ

本研究は日本の大学におけるEMIに必要な英語力を把握するニーズ分析を行った。インタビューデータを元にEMIのタスク一覧を作成し、それを元にタスクの頻度・難度を4段階で回答するアンケートを作成・実施した。その結果を元に、ほとんどの科目を英語で開講するEnglish-Taught Program(ETP)型と、選択必修科目にEMIが多いSemi-Structured(SS)型の各実施形態ごとに、EMIで経験する各タスクの頻度・難度を「高頻度・高難度」「高頻度・低難度」「低頻度・低難度」「低頻度・高難度」の4種類に分類した。ETP型とSS型のタスクの分類結果の共通点・相違点と自由回答記述データを比較・対照することで、優先して強化すべきEMIに必要な英語力や、逆に強化の優先順位の低いタスク、さらにETP型とSS型の違いを把握することができた。



■図4:本研究の全体像を示すジョイントディスプレイ

本研究の全体像を1つにまとめたものが図4のジョイントディスプレイである。図4が表すように、本研究全体で明らかになったことは、まずETP型とSS型の共通点として、EMI実施形態に関わらず高頻度・高難度のタスクは

- 文献を読む能力
- 読んだ内容をディスカッションやライティングとして表現する技能統合
- 学期末課題や卒論・修論などの長い論文の執筆であり、これが優先して強化すべきEMIで必要とされる英語力だと言える。一方、共通して低頻度・低難度のタスクは
- 日本語文献の活用
- 英語メールの作成
- 教室内でディスカッション中に友人に質問を投げかける行為

で、これらのタスクが出来るようになる訓練は優先順位が低いと言える。また、ETP型とSS型の相違点としては、ETP型では学生が提出した課題に関して教師からフィードバックを受ける機会が少ない一方で、SS型ではディスカッション中にクラスで話しづらい雰囲気を感じ取っている学生がいることが判明した。

6.2 本研究の意義

本研究の意義としては、目標状況分析とタスクを分析単位とする分析の組み合わせによるEMIのニーズ分析を実施できたことがある。実際にEMIを経験した参加者から得たインタビューデータを元に、EMIで経験する47個のタスク一覧を明らかにできた。その結果を元に作成・実施したアンケートを元に、SS型とETP型におけるEMIで経験するタスクの頻度・難度に関する情報を得ることができた。本研究は日本のある文脈における学部・大学院におけるニーズ分析ではあるものの、この結果は学習者から見たEMIの実施に関する知見を提供することができた。特に、日本に多いと考えられるSS型のEMIにおけるタスクの頻度・難度における情報、またそれを裏づける学習者が考える必要な能力・EMIのサポートに関する情報はEMIの実施に知見をもたらしたと言える。

本研究のもうひとつの意義はニーズ分析の方

法論的貢献ができたことにある。Serafini, Lake, and Long(2015)が推奨するニーズ分析の手順と、浦野(2005, 2014)が推奨する目標現状分析とタスクを組み合わせた手法は今後のEMIにおけるニーズ分析にも適用できると考えられる。さらに、GTxAを用いたアンケート作成の手法の有効性も示すことができた。インタビューデータを元にアンケートを作成し実施する作業は大変な労力を伴う。この手順の簡略化・高速化のために、テキストマイニングと人間の手動の分析を組み合わせるGTxAが、インタビューデータからタスクに関する情報を引き出してタスク一覧を作成するのに重要な役割を果たした。これらにより、本研究はニーズ分析における実施方法として方法論的示唆を提供できたと言える。

6.3 本研究の限界と今後の課題

本研究の限界を述べる。本研究はBrown and Iyobe(2014)の日本におけるEMIの6分類のうちETP型とSS型のみを対象とした。他の大学の別の種類のEMIプログラムではまた異なる結果が得られる可能性がある。また、本研究は学生を視点としたニーズ分析であるため、教員や他の利害関係者、教室で使用されている教材等は分析対象とはしなかった。これらのデータをさらに収集・分析し、本研究との結果と統合することでより包括的な「情報源と方法の組み合わせによるトライアンギュレーション」(Serafini, Lake, & Long, 2015)が達成されると考えられる。さらに、本研究は参加者間の比較を行う際に、ETP型とSS型との間の比較のみを実施した。その理由は資格試験の英語力スコアなど、参加者全員が共通して回答している背景情報が少なかったため、EMIの実施形態以外の要因を分析に加えることが不可能であったためである。今後、より参加者の背景情報がある場合は、例えば資格試験のスコアごとに難度の回答は変わるかどうかなどの調査が可能であると考えられる。

今後の研究において、異なる箇所におけるEMIのニーズ分析、特に本研究でデータを収集できなかった他の4種類のEMI環境においても研究を実施し、EMIの円滑な実施のためのデータを各地で収集・分析することが必要である。また

このニーズ分析の結果をカリキュラム開発・教材開発に繋げるための研究も必要である。例えば、本研究で重要とされた目標タスクに関してディスコースの収集（浦野, 2014）を実施し、それらを元にシラバスや教材開発を行うことが可能であり、それを EMI の準備コース等の開発へと適用できると考えられる。さらに、大学入学後に EMI を実施する学部・大学院の入学試験において、ニーズ分析の結果を元に各4技能外部試験の基準点の設定を行うなどの応用も必要であると考えられる。

謝辞

本研究を発表する貴重な機会を与えてくださいました公益財団法人日本英語検定協会と関係者の皆様、選考委員の先生方に深く御礼申し上げます。特に助言担当の和田稔先生からは貴重なアドバイスをいただき大変感謝しております。また、早稲田大学の澤木泰代先生からは研究計画のご助言からデータ収集のご協力まで様々な段階でお世話になりました。早稲田大学の原田哲男先生にもデータ収集の際にご協力をいただきました。さらに、早稲田大学の清田顕子さん、國橋さゆるさん、関佳奈さん、萩原ちはるさん、早崎綾さん、迎明香さんにはデータ収集のご協力から研究内容のフィードバックまで様々な点でお世話になりました。最後に、インタビュー・アンケートの回答に参加してくださった皆様に深く御礼申し上げます。

参考文献 (*は引用文献)

- Airey, J. (2016). EAP, EMI, or CLIL? In K. Hyland & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes*. Routledge (pp. 71-83). New York, NY:Routledge.
- Bachman, L. F., & Palmer, A. S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford University Press.
- Bradford, A. (2015). *Internationalization policy at the genba: Exploring the implementation of social science English-taught undergraduate degree programs in three Japanese universities* (Doctoral dissertation). The George Washington University, Washington.
- Bradford, A. (2016). Toward a typology of implementation challenges facing English-medium instruction in higher education: Evidence from Japan. *Journal of Studies in International Education*, 20(4), 339-356.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Brinton, D. M., & Snow, M. A. (2017). The evolving architecture of content-based instruction. In M. A. Snow & D. M. Brinton (Eds.), *The content-based classroom: New perspectives on integrating language and content* (pp. 2-20). Ann Arbor: University of Michigan Press.
- Brown, H. (2018). *Getting started with English-medium instruction in Japan: Key factors in program planning and implementation* (Doctoral dissertation). The University of Birmingham, Birmingham.
- Brown, H., & Iyobe, B. (2014). The growth of English medium instruction in Japan. *JALT2013 conference proceedings*, 9-19.
- Brown, J.D. (2009) Foreign and second language needs analysis. In M.H. Long & C.J. Doughty, (Eds.). *The Handbook of Language Teaching* (pp. 269-93). Oxford: Blackwell.
- Brown, J. D. (2016). *Introducing needs analysis and English for specific purposes*. Routledge.
- Chapple, J. (2015). Teaching in English is not necessarily the teaching of English. *International Education Studies*, 8(3), 1-13.
- Coyle, D., Hood, P., & Marsh, D. (2010). CLIL. Cambridge: Cambridge University Press.
- Dearden, J. (2014). *English as a medium of instruction-a growing global phenomenon*. British Council. https://www.britishcouncil.org/sites/default/files/e484_emi_-_cover_option_3_final_web.pdf 2020年5月28日閲覧
- Dudley-Evans, T. & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Fetters, M. D., & Freshwater, D. (2015). The 1+1=3 Integration Challenge. *Journal of Mixed Methods Research*, 9(2), 115-117.
- Harada, T. (2017). Developing a content-based English as a foreign language program: Needs analysis and curriculum design at the university level. In M. A. Snow & D. M. Brinton (Eds.), *The content-based classroom: New perspectives on integrating language and content* (pp. 37-52).
- Ann Arbor: University of Michigan Press.
- Hashimoto, K. (2013). 'English-only', but not a medium-of-instruction policy the Japanese way of internationalising education for both domestic and overseas students. *Current Issues in Language Planning*, 14(1), 16-33.
- Hess, J. M. (1968). Group interviewing. In R. L. King (Ed.), *New science of planning* (pp. 51-84). Chicago: American Marketing Association.
- 樋口耕一 (2004). 「テキスト型データの計量的分析—2つのアプローチの峻別と統合一」, 『理論と方法』, 19(1), 101-115.
- Hino, N. (2017). The significance of EMI for the learning of EIL in higher education: Four Cases from Japan. In B, Fenton-Smith, P, Humphreys, & I, Walkinshaw (Eds.), *English Medium Instruction in Higher Education in Asia-Pacific: From Policy to Pedagogy* (pp. 115-132). Springer.
- Inaba, M., & Kakai, H. (2019). Grounded Text Mining Approach: A Synergy between Grounded Theory and Text Mining Approaches. In A. Bryant & K. Charmaz (Eds.), *The SAGE Handbook of Current Developments in Grounded Theory* (pp. 332-351). Sage Publication.
- Kojima, N., & Yashima, T. (2017). Motivation in English medium instruction classrooms from the perspective of self-determination theory and the ideal self. *JACET Journal*, 61, 23-39.
- Kudo, S., Harada, T., Eguchi, M., Moriya, R., & Suzuki, S. (2017). Investigating English speaking anxiety in English-medium instruction. 『英語英文学叢誌』, 46(1), 7-23.
- 藏本真衣 (2017). 「TETE, EMI, CLILの定義:日本の英語教育に適した英語による英語教育を目指すためには」, *VERBA*, 40, 45-54.
- Krueger, R. A., & Casey, M. A. (2009). *Focus Groups: A Practical Guide for Applied Research* (4th ed.). Thousand Oaks, CA: Sage.
- Lambert, C. (2010). A task-based needs analysis: Putting principles into practice. *Language Teaching Research*, 14(1), 99-112.
- Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hylenstam & M. Pienemann (Eds.), *Modeling and assessing second language development* (pp. 77-99). Clevedon, Avon: Multilingual Matters.
- Long, M. H. (Ed.). (2005). *Second language needs analysis*. Cambridge University Press.
- Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, 51(1), 36-76.
- 文部科学省 (2011). 「グローバル人材育成推進会議中間まとめ」 http://www.kantei.go.jp/jp/singi/global/110622chukan_matome.pdf (2020年5月25日閲覧)
- 文部科学省 (2014). 「平成26年度「スーパーグローバル大学創成支援」採択構想の決定について」<https://warp.ndl.go.jp/info:ndljp/pid/11293659/www.mext.go.jp/> (2019年6月1日閲覧)
- Moncada-Comas, B., & Block, D. (2019). CLIL-isred EMI in practice: Issues arising. *The Language Learning Journal*,

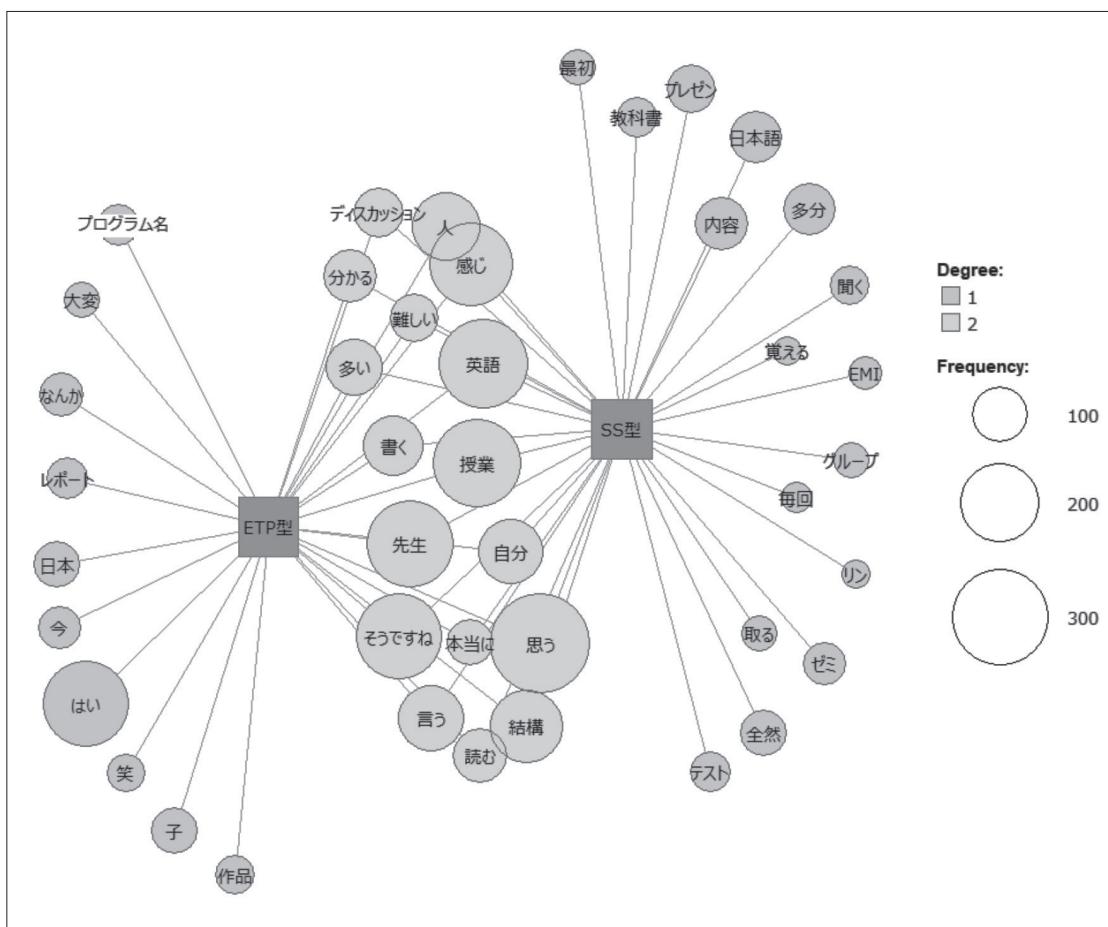
参考文献(*は引用文献)

- Online First.
- Morizumi, F. (2015). EMI in Japan: Current status and its implications. *Educational Studies*, 57, 119-128.
- Murata, K. (2019). The realities of the use of English in the globalised world and the teaching of English: A discrepancy? (Invited papers by lecturers for JACET 44th Summer Seminar 2017, Tokyo). *JACET Journal*, 63, 7-26.
- 村田久美子, 飯野公一, 小中原麻友 (2017).「EMI(英語を媒介とする授業)における「共通語としての英語」の使用の現状把握と意識調査、および英語教育への提言」.『早稲田教育評論』, 31(1), 21-38.
- Ota, H. (2018). Internationalization of higher education Global trends and Japan's challenges. *Educational Studies in Japan: International Yearbook*, 12, 91-105.
- Ortega, L. (2015). Researching CLIL and TBLT interfaces. *System*, 54, 103-109.
- Poole, G. (2017). Administrative impediments: How bureaucratic practices obstruct the implementation of English-Taught Programs in Japan In A. Bradford & B. Howard (Eds.), *English-Medium Instruction in Japanese Higher Education: Policy, Challenges and Outcomes* (pp. 91-107). Multilingual Matters.
- Richards, J. (2001). *Curriculum development in language teaching*. Cambridge: Cambridge University Press.
- Rose, H., & McKinley, J. (2018). Japan's English-medium instruction initiatives and the globalization of higher education. *Higher Education*, 75(1), 111-127.
- Rosenfeld, M., Leung, S., & Oltman, P. (2001). Identifying the reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels. TOEFL monograph series MS-21. Princeton, NJ: Educational Testing Service.
- Sawaki, Y. (2017). University faculty members' perspectives on English language demands in content courses and a reform of university entrance examinations in Japan: A needs analysis. *Language Testing in Asia*, 7(13), 1-16.
- Serafini, E. J., Lake, J. B., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes*, 40, 11-26.
- Serafini, E. J., & Torres, J. (2015). The utility of needs analysis for nondomain expert instructors in designing task - based Spanish for the professions curricula. *Foreign Language Annals*, 48(3), 447-472.
- Shimauchi, S. (2018). English-Medium Instruction in the internationalization of higher education in Japan: Rationales and issues. *Educational Studies in Japan: International Yearbook*, 12, 77-90.
- Snow, M. A. (2014). Content-based and immersion models of second/foreign language teaching. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed., pp. 438-454). Boston, MA: National Geographic Learning/Heinle Cengage Learning.
- Suzuki, S., Harada, T., Eguchi, M., Kudo, S., & Moriya, R. (2018). Students' perspectives on the role of English-medium instruction in English learning: A case study. *The Bulletin of the Graduate School of Education of Waseda University, Separate Volume*, 26(1), 1-18.
- Taguchi, N. (2014). Pragmatic socialization in an English-medium university in Japan. *IRAL*, 52(2), 157-181.
- Tsuneyoshi, R. (2005). Internationalization strategies in Japan: The dilemmas and possibilities of study abroad programs using English. *Journal of Research in International Education*, 4(1), 65-86.
- Uchihara, T., & Harada, T. (2018). Roles of vocabulary knowledge for success in English - medium instruction: Self - perceptions and academic outcomes of Japanese undergraduates. *TESOL Quarterly*, 52(3), 564-587.
- 浦野研 (2005, 10月)「ニーズ分析の方法論:Who, What, & How?」ESP 北海道第1回公開研究会, 道都大学研究セントター https://www.urano-ken.com/research/ESP_Hokkaido_handout.pdf (2020年5月25日閲覧)
- 浦野研 (2014, 12月)「ESP カリキュラム開発のためのタスクに基づいたニーズ分析」. 外国語教育メディア学会関西支部メソドロジー研究部会2014年度 <https://www.slideshare.net/uranoken/methoken2014> (2020年5月25日閲覧)
- ヴォーン, S., シューム, J. S., & シナグブ, J. (1999). 井上理監訳・田部井潤・柴原宣幸(訳)『グループ・インタビューの技法』. 東京: 慶應義塾大学出版会
- West, R. (1994). Needs analysis in language teaching. *Language Teaching*, 27, 1-19.

資料1：インタビューデータのテキストマイニング分析結果(抽出語リストの一部).....

名詞		サ変名詞		形容動詞	
英語	422	授業	332	大変	48
先生	410	ディスカッション	145	普通	42
感じ	326	レポート	70	確か	33
自分	201	テスト	57	好き	32
日本語	172	話	57	必要	23
内容	122	理解	48	大事	22
ゼミ	112	発表	43	大丈夫	17
グループ	75	意見	41	得意	16
文学	72	発言	40	完全	13
最初	61	説明	38	自由	13
教科書	60	一緒	37	単純	12
課題	57	意味	36	たしか	10
作品	54	質問	32	アカデミック	9
基本	51	留学	29	苦手	9
プレゼンテーション	46	フィードバック	28	結構	9
言語	46	選択	28	微妙	9
専門	46	研究	27	簡単	7
論文	42	勉強	26	変	7
最後	39	レクチャー	24	ネイティブ	6
一つ	36	演習	24	メイン	6
必修	34	教育	21	楽	6
学部	32	翻訳	21	嫌い	6
日本人	32	経験	20	正直	6
ページ	31	講義	18	無理	6
ペーパー	31	準備	18	あんまり	5
個人	29	分析	18	いや	5
トピック	28	メール	16	同じ	5
リン	26	イメージ	14	だめ	4
興味	25	スライド	14	めちゃめちゃ	4
バイ	24	プリント	14	コンスタント	4
学生	24	担当	14	ダメ	4
テーマ	23	発音	14	自然	4
海外	23	意識	13	不安	4

資料2: インタビューデータのテキストマイニング分析結果(共起ネットワーク図).....



資料3: ETP型のアンケート結果の量的データ(頻度・難度のパーセンテージと平均値).....

タスク一覧	頻度					難度					
	4	3	2	1	経験なし	4	3	2	1	経験なし	平均値
1. 専門科目の内容をインターネットで英語で調べる	0.42	0.50	0.08	0.00	0.00	3.33	0.08	0.38	0.33	0.21	0.00 2.33
2. 英語で書かれた専門の教科書を読む際に分からない単語を調べる	0.46	0.42	0.13	0.00	0.00	3.33	0.04	0.21	0.46	0.29	0.00 2.00
3. 英語で書かれた専門の教科書の指定された章を読む	0.71	0.21	0.08	0.00	0.00	3.63	0.04	0.46	0.42	0.08	0.00 2.46
4. 英語で書かれた専門の教科書の概要・大意を把握する	0.63	0.25	0.13	0.00	0.00	3.50	0.17	0.21	0.42	0.21	0.00 2.33
5. 英語で書かれた専門の教科書の中から特定の情報を探し出す	0.50	0.33	0.13	0.04	0.00	3.29	0.13	0.38	0.38	0.13	0.00 2.50
6. 専門内容を理解するために英語で書かれた専門の教科書をじっくり読む	0.33	0.29	0.29	0.08	0.00	2.88	0.25	0.29	0.38	0.08	0.00 2.71
7. 複数の英語で書かれた文献を読んで内容を比較する	0.21	0.33	0.42	0.04	0.00	2.71	0.25	0.33	0.38	0.04	0.00 2.79
8. 英語で書かれたリーディング課題を和訳する	0.13	0.00	0.17	0.58	0.13	1.62	0.08	0.17	0.42	0.21	0.13 2.14
9. 専門科目の内容について質問するために先生に英語でメールを送る	0.21	0.33	0.25	0.17	0.04	2.61	0.04	0.17	0.33	0.42	0.04 1.83
10. 専門授業に関して交渉するために先生に英語でメールを送る	0.04	0.08	0.46	0.29	0.13	1.86	0.04	0.29	0.25	0.33	0.08 2.05
11. 専門用語を英語で定義し例を挙げて書く	0.08	0.33	0.29	0.21	0.08	2.32	0.04	0.58	0.21	0.13	0.04 2.57
12. 先生から専門内容についての説明を英語で聞く	0.79	0.13	0.08	0.00	0.00	3.71	0.04	0.33	0.42	0.21	0.00 2.21
13. 英語で書かれた専門の教科書の重要な点を英語で書く	0.46	0.33	0.13	0.08	0.00	3.17	0.00	0.25	0.50	0.25	0.00 2.00
14. 先生からの専門内容についての説明をノートに英語で書く	0.71	0.25	0.04	0.00	0.00	3.67	0.00	0.21	0.33	0.46	0.00 1.75
15. 先生に専門内容について英語で質問する	0.33	0.33	0.13	0.21	0.00	2.79	0.13	0.29	0.38	0.21	0.00 2.33
16. 先生の意見に対して先生に英語で反論する	0.08	0.25	0.17	0.38	0.13	2.05	0.25	0.38	0.21	0.04	0.13 2.95
17. 専門内容に関して英語教材(音声・映像)を視聴する	0.63	0.13	0.08	0.13	0.04	3.30	0.04	0.17	0.42	0.33	0.04 1.91
18. 授業で使用される、英語で書かれた資料(ハンドアウト・スライド)の内容を理解する	0.71	0.29	0.00	0.00	0.00	3.71	0.00	0.13	0.42	0.46	0.00 1.67
19. 他の学生が書いた英語の作品を英語で口頭でフィードバックする	0.25	0.21	0.21	0.08	0.25	2.83	0.17	0.17	0.29	0.13	0.25 2.50
20. 日本語で書かれた資料について英語で口頭で質問する	0.04	0.17	0.29	0.25	0.25	2.00	0.04	0.08	0.42	0.21	0.25 1.94
21. 教場テストでエッセイを英語で書く	0.38	0.50	0.13	0.00	0.00	3.25	0.04	0.50	0.21	0.25	0.00 2.33
22. 英語で書かれた専門の教科書で読んだ内容についてディスカッションをする	0.50	0.25	0.25	0.00	0.00	3.25	0.21	0.42	0.29	0.08	0.00 2.75
23. 与えられた質問に対しペアで英語でディスカッションする	0.42	0.42	0.08	0.04	0.04	3.26	0.04	0.25	0.50	0.21	0.00 2.13
24. ディスカッション中に専門内容について英語で説明する	0.21	0.33	0.38	0.04	0.04	2.74	0.04	0.38	0.46	0.08	0.04 2.39
25. ディスカッション中に専門内容について他の学生に英語で質問をする	0.25	0.42	0.25	0.08	0.00	2.83	0.04	0.29	0.42	0.25	0.00 2.13
26. ディスカッション中に他の学生に発言を英語で促す	0.13	0.25	0.29	0.21	0.13	2.33	0.17	0.13	0.21	0.33	0.17 2.15
27. グループ代表としてグループディスカッションの内容をクラス全体に紹介する	0.04	0.25	0.29	0.25	0.17	2.10	0.25	0.38	0.21	0.04	0.13 2.95
28. 専門内容について英語でクラス全体でディスカッションをする	0.38	0.25	0.13	0.21	0.04	2.83	0.33	0.33	0.25	0.04	0.04 3.00
29. 専門科目の内容についてプレゼンテーションのスライドを英語で書く	0.75	0.21	0.00	0.04	0.00	3.67	0.00	0.25	0.33	0.42	0.00 1.83
30. 英語で書かれた専門の教科書で読んだ内容のまとめをプレゼンテーションする	0.50	0.42	0.04	0.00	0.04	3.48	0.08	0.33	0.33	0.21	0.04 2.30
31. 専門科目の内容についてグループで英語でプレゼンテーションをする	0.50	0.33	0.08	0.04	0.04	3.35	0.04	0.33	0.42	0.17	0.04 2.26
32. プrezentationでディスカッションクエスチョンを英語で投げかける	0.38	0.21	0.33	0.04	0.04	2.96	0.04	0.21	0.42	0.25	0.08 2.05
33. 自分の発表した内容について先生や学生からの意見を英語で聞く	0.42	0.38	0.17	0.04	0.00	3.17	0.08	0.25	0.25	0.42	0.00 2.00
34. 他の学生から専門内容についての説明を英語で聞く	0.50	0.33	0.08	0.04	0.04	3.35	0.04	0.21	0.38	0.33	0.04 1.96
35. プrezenterが投げかけたディスカッションクエスチョンに回答する	0.42	0.25	0.13	0.13	0.08	3.05	0.08	0.29	0.21	0.29	0.13 2.19
36. プrezentationの内容について口頭でフィードバックする	0.25	0.17	0.38	0.08	0.13	2.67	0.08	0.42	0.21	0.13	0.17 2.55
37. 学生のprerezentationについて評価シート等でフィードバックを英語で書く	0.17	0.21	0.42	0.13	0.08	2.45	0.00	0.17	0.38	0.33	0.13 1.81
38. 日本語で書かれた資料について英語で書く	0.21	0.08	0.29	0.25	0.17	2.30	0.04	0.08	0.50	0.21	0.17 1.95
39. 英語で書かれた専門の教科書の要約を英語で書く	0.25	0.21	0.29	0.21	0.04	2.52	0.04	0.46	0.25	0.21	0.04 2.35
40. 英語で書かれた専門の教科書で読んだ内容について自分の考察を英語で書く	0.46	0.21	0.08	0.17	0.08	3.05	0.13	0.58	0.08	0.13	0.08 2.77
41. レポートで引用するために参考文献を英語で読む	0.67	0.29	0.00	0.04	0.00	3.56	0.13	0.33	0.29	0.25	0.00 2.33
42. 専門科目の内容についてレポートを英語で書く	0.83	0.13	0.04	0.00	0.00	3.79	0.13	0.42	0.25	0.21	0.00 2.46
43. 専門科目の内容についてグループで分担して英語でレポートを書く	0.08	0.21	0.29	0.21	0.21	2.21	0.04	0.33	0.25	0.17	0.21 2.32
44. 英語のライティング課題を他の学生からフィードバックを受けて修正して提出する	0.17	0.00	0.25	0.29	0.29	2.06	0.04	0.08	0.25	0.33	0.29 1.76
45. 専門内容の知識を利用して英語で創作をする(例:小説・詩・演劇, etc...)	0.08	0.00	0.13	0.29	0.50	1.75	0.21	0.00	0.17	0.13	0.50 2.58
46. 専門科目の内容について学期末レポートを英語で書く	0.83	0.17	0.00	0.00	0.00	3.83	0.08	0.42	0.38	0.13	0.00 2.46
47. 学位論文(卒業論文や修士論文)を英語で書く	0.50	0.00	0.00	0.17	0.33	3.25	0.29	0.25	0.04	0.08	0.33 3.13

(注)頻度・難度共に「4=高い」、「3=やや高い」、「2 = やや低い」、「1=低い」を表す。平均値は「経験なし」を0とみなした各タスクごとの平均値を表す。

資料4: SS型のアンケート結果の量的データ(頻度・難度のパーセンテージと平均値)

タスク一覧	頻度					難度						
	4	3	2	1	経験なし	平均値	4	3	2	1	経験なし	平均値
1. 専門科目の内容をインターネットで英語で調べる	0.36	0.24	0.27	0.12	0.00	2.85	0.12	0.42	0.36	0.09	0.00	2.58
2. 英語で書かれた専門の教科書を読む際に分からぬ単語を調べる	0.52	0.30	0.12	0.06	0.00	3.27	0.06	0.30	0.39	0.24	0.00	2.18
3. 英語で書かれた専門の教科書の指定された章を読む	0.76	0.18	0.00	0.06	0.00	3.64	0.27	0.39	0.33	0.00	0.00	2.94
4. 英語で書かれた専門の教科書の概要・大意を把握する	0.55	0.30	0.12	0.03	0.00	3.36	0.24	0.45	0.30	0.00	0.00	2.94
5. 英語で書かれた専門の教科書の中から特定の情報を探し出す	0.39	0.45	0.09	0.06	0.00	3.18	0.30	0.30	0.30	0.09	0.00	2.82
6. 専門内容を理解するために英語で書かれた専門の教科書をじっくり読む	0.36	0.36	0.15	0.12	0.00	2.97	0.52	0.33	0.12	0.03	0.00	3.33
7. 複数の英語で書かれた文献を読んで内容を比較する	0.24	0.21	0.30	0.12	0.12	2.66	0.45	0.21	0.21	0.00	0.12	3.28
8. 英語で書かれたリーディング課題を和訳する	0.09	0.06	0.39	0.30	0.15	1.93	0.18	0.27	0.27	0.12	0.15	2.61
9. 専門科目の内容について質問するために先生に英語でメールを送る	0.06	0.12	0.15	0.36	0.30	1.83	0.12	0.21	0.24	0.12	0.30	2.48
10. 専門授業に関して交渉するために先生に英語でメールを送る	0.03	0.09	0.27	0.21	0.39	1.90	0.12	0.15	0.30	0.06	0.36	2.52
11. 専門用語を英語で定義し例を挙げて書く	0.21	0.30	0.24	0.18	0.06	2.58	0.33	0.30	0.27	0.03	0.06	3.00
12. 先生から専門内容についての説明を英語で聞く	0.67	0.12	0.09	0.12	0.00	3.33	0.21	0.33	0.36	0.09	0.00	2.67
13. 英語で書かれた専門の教科書の重要な点を英語で書く	0.24	0.30	0.15	0.24	0.06	2.58	0.24	0.36	0.30	0.03	0.06	2.87
14. 先生からの専門内容についての説明をノートに英語で書く	0.18	0.42	0.18	0.09	0.12	2.79	0.18	0.30	0.24	0.15	0.12	2.59
15. 先生に専門内容について英語で質問する	0.30	0.27	0.18	0.15	0.09	2.80	0.21	0.39	0.24	0.06	0.09	2.83
16. 先生の意見に対して先生に英語で反論する	0.06	0.12	0.39	0.21	0.21	2.04	0.45	0.15	0.12	0.06	0.21	3.27
17. 専門内容に関して英語教材(音声・映像)を視聴する	0.15	0.30	0.33	0.18	0.03	2.44	0.15	0.36	0.24	0.21	0.03	2.47
18. 授業で使用される、英語で書かれた資料(ハンドアウト・スライド)の内容を理解する	0.76	0.18	0.03	0.03	0.00	3.67	0.06	0.24	0.64	0.06	0.00	2.30
19. 他の学生が書いた英語の作品を英語で口頭でフィードバックする	0.15	0.12	0.33	0.15	0.24	2.36	0.18	0.33	0.21	0.00	0.27	2.96
20. 日本語で書かれた資料について英語で口頭で質問する	0.03	0.12	0.27	0.27	0.30	1.87	0.12	0.24	0.18	0.12	0.33	2.55
21. 教場テストでエッセイを英語で書く	0.18	0.18	0.24	0.21	0.18	2.41	0.24	0.33	0.18	0.06	0.18	2.93
22. 英語で書かれた専門の教科書で読んだ内容についてディスカッションをする	0.55	0.27	0.12	0.06	0.00	3.30	0.39	0.42	0.18	0.00	0.00	3.21
23. 与えられた質問に対しペアで英語でディスカッションする	0.45	0.24	0.21	0.09	0.00	3.06	0.30	0.36	0.24	0.09	0.00	2.88
24. ディスカッション中に専門内容について英語で説明する	0.27	0.33	0.24	0.12	0.03	2.78	0.39	0.30	0.24	0.03	0.03	3.09
25. ディスカッション中に専門内容について他の学生に英語で質問をする	0.21	0.33	0.30	0.12	0.03	2.66	0.24	0.27	0.39	0.06	0.03	2.72
26. ディスカッション中に他の学生に発言を英語で促す	0.09	0.24	0.30	0.21	0.15	2.25	0.18	0.15	0.36	0.15	0.15	2.43
27. グループ代表としてグループディスカッションの内容をクラス全体に紹介する	0.15	0.33	0.18	0.24	0.09	2.43	0.30	0.30	0.27	0.03	0.09	2.97
28. 専門内容について英語でクラス全体でディスカッションをする	0.33	0.30	0.15	0.18	0.03	2.81	0.48	0.27	0.18	0.03	0.03	3.25
29. 専門科目の内容についてプレゼンテーションのスライドを英語で書く	0.48	0.30	0.09	0.06	0.06	3.29	0.30	0.36	0.27	0.00	0.06	3.03
30. 英語で書かれた専門の教科書で読んだ内容のまとめをプレゼンテーションする	0.42	0.33	0.09	0.09	0.06	3.16	0.24	0.55	0.12	0.03	0.06	3.06
31. 専門科目の内容についてグループで英語でプレゼンテーションをする	0.18	0.27	0.30	0.12	0.12	2.59	0.27	0.36	0.21	0.03	0.12	3.00
32. プrezentationでディスカッションクエスチョンを英語で投げかける	0.33	0.33	0.12	0.06	0.15	3.11	0.15	0.42	0.18	0.09	0.15	2.75
33. 自分の発表した内容について先生や学生からの意見を英語で聞く	0.36	0.33	0.12	0.12	0.06	3.00	0.21	0.21	0.45	0.06	0.06	2.61
34. 他の学生から専門内容についての説明を英語で聞く	0.42	0.27	0.15	0.09	0.06	3.10	0.12	0.30	0.42	0.09	0.06	2.48
35. プrezenterが投げかけたディスカッションクエスチョンに関してグループで話し合う	0.42	0.42	0.06	0.03	0.06	3.32	0.12	0.42	0.30	0.09	0.06	2.61
36. プrezentationの内容について口頭でフィードバックをする	0.09	0.30	0.33	0.15	0.12	2.38	0.27	0.24	0.27	0.09	0.12	2.79
37. 学生のprerezentationについて評価シート等でフィードバックを英語で書く	0.33	0.18	0.21	0.12	0.15	2.86	0.06	0.21	0.36	0.21	0.15	2.14
38. 日本語で書かれた資料について英語で書く	0.03	0.12	0.33	0.27	0.24	1.88	0.12	0.24	0.21	0.15	0.27	2.46
39. 英語で書かれた専門の教科書の要約を英語で書く	0.15	0.33	0.27	0.15	0.09	2.53	0.24	0.30	0.30	0.06	0.09	2.80
40. 英語で書かれた専門の教科書で読んだ内容について自分の考察を英語で書く	0.21	0.30	0.33	0.12	0.03	2.63	0.33	0.42	0.18	0.03	0.03	3.09
41. レポートで引用するために参考文献を英語で読む	0.58	0.33	0.03	0.03	0.03	3.50	0.24	0.39	0.24	0.09	0.03	2.81
42. 専門科目の内容についてレポートを英語で書く	0.55	0.30	0.03	0.09	0.03	3.34	0.33	0.36	0.24	0.03	0.03	3.03
43. 専門科目の内容についてグループで分担して英語でレポートを書く	0.09	0.30	0.27	0.15	0.18	2.41	0.21	0.39	0.18	0.03	0.18	2.96
44. 英語のライティング課題を他の学生からフィードバックを受けて修正して提出する	0.21	0.06	0.45	0.12	0.15	2.43	0.21	0.30	0.27	0.06	0.15	2.79
45. 専門内容の知識を利用して英語で創作をする(例:小説・詩・演劇, etc...)	0.00	0.00	0.06	0.36	0.58	1.14	0.21	0.12	0.06	0.03	0.58	3.21
46. 専門科目の内容について学期末レポートを英語で書く	0.55	0.27	0.06	0.09	0.03	3.31	0.42	0.36	0.18	0.00	0.03	3.25
47. 学位論文(卒業論文や修士論文)を英語で書く	0.42	0.24	0.03	0.06	0.24	3.36	0.55	0.09	0.15	0.00	0.21	3.50

(注) 頻度・難度共に「4=高い」、「3=やや高い」、「2=やや低い」、「1=低い」を表す。平均値は「経験なし」を0とみなした各タスクごとの平均値を表す。

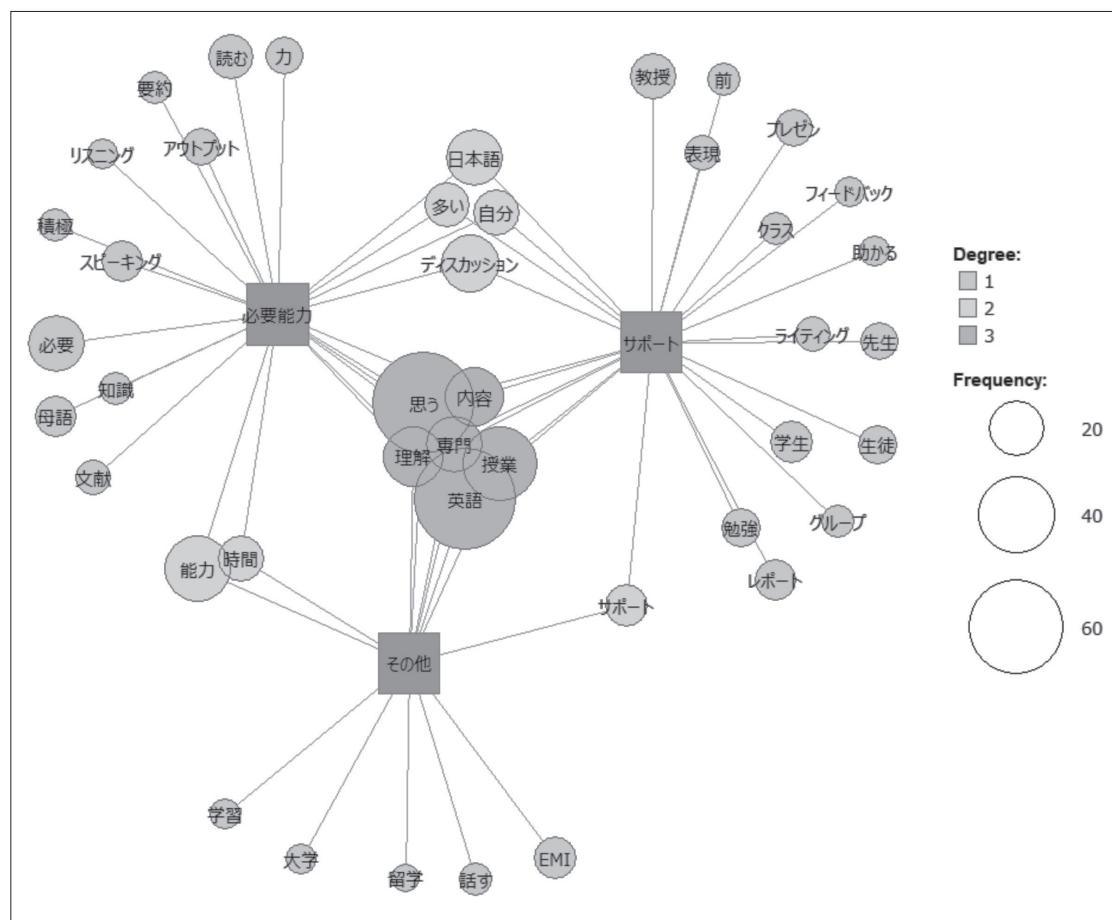
資料5：EMIに必要な能力に対するテーマ分析結果(抜粋含む)

サブカテゴリー	内訳	詳細(抜粋)
技能・能力(30)	リーディング(12)	「文献を素早く読み内容を取る能力(ETP)」「文章処理のスピード(ETP)」「パラグラフごとの要約をつかみながら文章を読む能力。ある程度の速度で多くの分量を読む必要があったため。(ETP)」
	スピーキング(7)	「基本的なスピーキング能力をしっかりと高校時代にも身に付けていなければ、いきなりEMIの授業でディスカッションをすることは難しい(SS)」「ディスカッションがとても多かったので、伝えたいことを伝わるように説明するのは難しかった(ETP)」
	リスニング(3)	「教員やクラスメイトの英語での発言をちゃんと理解するため、リスニング能力が必要(SS)」「コミュニケーション能力も大切だと思いました。ディスカッションがとても多かったので、伝えたいことを伝わるように説明するのは難しかったです(ETP)」
	ライティング(2)	「ライティングの能力がより高ければ、レポート等の課題がより短時間で質良く仕上げられた(SS)」「ライティング(特に要約)の能力(SS)」
	その他(6)	「要約力。英語文献からどこが大切なかを抽出したり、要約するのが苦手だった(SS)」「英語を日本語に訳さないで英語におき換える練習、能力が必要(SS)」「オンライン上やメールなどでレポートをフィードバックしてもらえるシステムがあれば有り難い(ETP)」
知識・理解(11)	日本語による理解(6)	「日本語で専門科目について基礎的な知識があれば英語での説明も飲み込みやすかった(SS)」
	専門知識(5)	「内容に関する知識に詳しいこと。話し合いなどがスムーズになる(SS)」
授業内の発言(8)	アウトプット(4)	「在学中受けていた授業では、インプットの場は多かったが、それをアウトプットする機会が少なかった(SS)」
	発信力(4)	「そもそも発言する力?発言しやすい環境?が必要かと思いました(SS)」
情意面(9)	情意面(9)	「ディスカッションに参加する積極性と自信が必要(ETP)」「拙いながらも英語で発言する勇気(ETP)」「母国語で授業を受けるよりもしっかりと聞かないと理解することが出来ないので継続した集中力が必要(SS)」

資料6：EMIに必要なサポートに対するテーマ分析結果(抜粋含む)

サブカテゴリー	内訳	詳細(抜粋)
アカデミックスキル(19)	全般(6)	「初期の段階で専門書の効率的な読み方などが学べる授業があると良いと思います(SS)」
	スピーキング(2)	「英語でのプレゼンやディスカッションは躊躇してしまうので、人前で躊躇せずに英語で発言できるだけのスピーキング能力をつけたい(SS)」
	ライティング(6)	「英語でのレポートの書き方 最初の頃は書式も分からず大変だったので(SS)」
	使用機会(5)	「ディスカッション練習の場(ETP)」「英語ディスカッションを練習するクラス(ETP)」
専門知識・内容(12)	内容(4)	「専門用語について復習の時間(SS)」「発信した内容や言葉のチョイスで言いたいことが伝えられたか、確認するプロセスがあるとよい(SS)」
	日本語による補助(8)	「入門科目の前に母語での入門科目を設置する(SS)」「現在は、多くの情報が日本語でも載っているので、outputを英語ではるのは大切だが多くの情報をinputする際は日本語でもいいと思う(ETP)」
理解・確認(9)	フィードバック(4)	「日本の大学はレポートなどにフィードバックがないのが多く、学生一人ひとりのきめ細やかなフィードバックがあれば、自分の弱点も見えて勉強しやすくなるかなと留学して思った(ETP)」「レポートを書いてもほとんどどのクラス(特に大教室)で教授からのフィードバックがなかったため、自分のレポートのどこが良かったのか、どこを改善すればいいのか全くわからなかった(ETP)」
	TAやピアによるサポート(5)	「留学先にいたATは教授よりも距離が近いため質問しやすかった。ATの存在は授業内容理解に大きく寄与したと思う(ETP)」「言語パートナーをつける(ETP)」
学習環境(24)	授業進行(13)	「毎回のディスカッションの内容が似たものにならなかった。同じ質問でも「○○点から」などポイントの指定があれば、また違った角度からディスカッションができる、より深く議論ができるように思う。(SS)」「授業中にグループ分けをして、各グループに発表させる授業では、皆が比較的の発言できていた(ETP)」
	雰囲気・設備(7)	「質問する雰囲気が重たい、しづらい(SS)」「教授と生徒が双方向で意見を言える環境があると、わからない部分を聞けたり自分も参加しているという意識ももってて理解力も上がると思う(SS)」
	時間・精神的ゆとり(4)	「授業に積極的に参加できるか否かは、単純な英語能力だけによるものではないと思う。学生の心理(パーソナリティ)に関係する面も多くあると思う。その面も踏まえて、学生の精神的なケアも必要となると思う。(SS)」

資料7：自由回答記述データのテキストマイニングの出力結果



C 調査部門・報告Ⅱ・英語教育関連の調査・アンケートの実施と分析

日本人大学生の英語学習における 自己調整学習能力尺度の開発:英語資格試験に 向けた自主学習に焦点を当てて

研究者:東京都／立教大学大学院 在籍 福田 晶子

《研究助言者:吉田 研作》

概要

本研究の目的は、大学生の英語資格試験に向けた自主学習に焦点を当て、自己調整学習能力を測るための尺度を開発し、その妥当性を検証することにある。自己調整学習の理論的背景をもとに情意尺度、動機づけ調整方略尺度、学習方略尺度の3点からなる自己調整学習能力尺度を作成した。先行研究とインタビュー調査を通して計122項目を作成したのち、5件法により356人の日本人大学生から回答を得た。それぞれの尺度に対し探索的因子分析を実施し、情意尺度で3因子、動機づけ調整方略尺度で2因子、学習方略尺度で5因子の因子構造が得られた。この尺度の信頼性と妥当性を確認するために、確証的因子分析・相関分析・重回帰分析・パス解析を行ったところ、いずれも、これら3つの尺度が自己調整学習方略尺度として一定の水準に達していることが明らかになった。

1 はじめに

日本のような英語を外国語として学習する環境下で英語を習得するためには、学校教育で英語に触れるほかに、自主的に学習機会を増やすことが必要だとわれている(廣森, 2015)。この自発的な学習過程を理論化したものが、自己調整学習(Zimmerman, 2000)である。自主学習において自己調整学習は重要な役割を果たすと考えられるが、教室環境下を離れた英語学習にまつわる研

究は少なく、中でも、自主学習における自己調整学習の研究は、ほんないと言ってよい。一方で、大学生の自主学習と一口に言っても、どのような目的で自主学習を行うかによって、その作業方法や内容は大きく異なる。そこで本研究では、自主学習の目的を英語資格試験の受験と定め、それに向けた自主学習における自己調整学習能力を測定するための尺度を作成した。これは、宿題やレポート課題と異なり、教員や授業の成績といった制約と直接的な関係がない点で、学習者の自律性、すなわち自己調整学習能力が求められる活動だと判断したからである。大学生は英語資格試験に向けて、どのような自己調整学習を行うのかを、尺度開発を通して理解する。

2

研究の背景と目的

2.1 自己調整学習とは

日本の英語教育研究において、自己調整学習は近年注目を浴びているトピックである(英語教育, 2019, 2020)。自己調整学習は「学習者が目標達成に向けて、自らの認知・動機・行動・情意を活性化させ、維持しながら学習を進める過程」と定義されている(Schunk & Greene, 2018)。1980年代から本格的に行われてきた自己調整学習研究の成果の一つとして、自己調整学習能力は学業成績と正の相関関係にあることがわかっている

る(Pintrich & De Groot, 1990; Zimmerman & Martinez-Pons, 1990)。これは、第二言語習得研究においても同様の知見が得られており、より自己調整学習ができる学習者ほど、言語学習に邁進し、結果的により良い成果に到達するということが明らかになっている(Seker, 2016)。

自己調整学習を行う上で重要な3要素として、Zimmerman(1989)は「自己効力感」「目標関与」「自己調整学習方略」を挙げている。自己効力感とは課せられた課題や活動に対する自分の能力を信じること(Mills, 2014)であり、簡単に言うと、学習に対して学習者が自分ならできると思える自信である。目標関与については、自己調整学習を進めるためには目標の存在が不可欠だと言われており、自分自身に関わりのある目標を立てることが自己調整学習を促す始発点と考えられている。そして、自己調整学習方略とは主に「メタ認知方略」のことを指している。これは、学習者自身が、学習過程の段階で計画を立て、学習内容に対する理解度の把握や進捗を管理し、自己評価や問題対処をしながら学習する能力であり、自己調整学習を実施するための方略である(伊藤, 2009)。つまり、自己調整学習能力のある学習者は、自己に対する期待や信念を持ち、自分自身で学習内容や到達目標を決めることができ、それに向かって学習する上で、やる気を奮い立たせたり、自分の理解度を把握したり、適切な学習方法を模索したりしながら、継続して学習することができると考えられている(Zimmerman, 2001)。このように、自律的な学習者の特徴を理解する上でも、自律的な学習を支援する上でも、自己調整学習という枠組みは有用だとされている。

ただし、自己調整学習の概念と下位要素についてのあらゆる検討は、基本的に教室環境下での学習形態を基準として行われてきた。Schunk & Greene(2018)は、検証るべき文脈の拡充が今後の自己調整学習研究の課題点であると述べている。特に、学校環境下を離れた学習環境(out-of-school contexts)での自己調整学習の解明が求められている。

2.2 第二言語習得研究における自己調整学習の尺度開発

第二言語習得研究に自己調整学習という概念

が導入されたのは2000年代だが、それまでは学習方略研究や動機づけ研究といったように、学習者要因に対して別個のアプローチがとられてきた。そのような中で、学習方略研究の定義の曖昧性に対する批判や、学習者要因を個別に検証するのではなく包括的に理解する必要性があるといった主張により、自己調整学習に対する関心が高まっていった(Dörnyei, 2005)。同時に、どのような自己調整学習能力が第二言語能力と関連しているかについて質問紙が作成され、検証されてきた。例えば、Teng, Dörnyei, & Schmitt(2006)は語彙学習における自己調整学習能力を測る質問紙を提案し、さまざまな文化的背景(例えば、Mizumoto & Takeuchi, 2012)から追試されている。

特定の技能ではなく、第二言語習得全般について扱った質問紙も、数が少ないものの存在する。例えば、Habok & Magyar(2018)は、教室での外国語活動における自己調整学習能力の尺度(Self-regulated foreign language learning strategy questionnaire, SRFLSQ)を開発している。ハンガリーの中学生を対象とし、外国語の授業における学習態度を自己調整学習の枠組みで検討している。Wang, Schwab, Fenn, & Chang(2013)は、中国人大学生とドイツ人大学生の英語学習を比較するために、自己調整学習尺度(The questionnaire of English self-regulated learning strategies, QESRLS)を開発し、自己効力感との関係性を検証している。この2つの質問紙は、Oxford(2011, 2017)による第二言語学習方略に特化した自己調整モデル(Strategic Self-regulation, S²R)を理論的背景としており、一定の信頼性と妥当性を持つことが確認されている。しかし、SRFLSQの場合、中学生の英語授業における態度の測定を試みている点で、大学生の英語学習とは文脈が大きく乖離していると言えよう。さらに、いずれも教室環境下における英語学習を想定した尺度であり、英語の授業内での行動について自己調整学習の観点から検討されている。つまり、本研究のねらいである自主学習環境に適した自己調整学習能力尺度はいまだ開発されていない。

2.3 英語自主学習と自己調整学習

本研究で自主学習に着目する理由は、外国语として英語を学ぶ(English as a Foreign Language, EFL)環境下で英語熟達度を高めるには、自主学習が重要な鍵を握っているからである。英語を第二言語とする(English as a Second Language, ESL)学習者は、日常生活を英語で行う必要性が付きまとつたため、必然的に第二言語(英語)の使用機会(インプット・アウトプット)が増える。一方で、日本のようなEFL環境下では、時間的制約は自然に打破できるものではない。例えば、廣森(2015)は、教室環境下における外国语教育で確保されている授業時数の試算をもとに、中学・高校・大学を合わせた平均的な学習時間と小学校で教科として英語を学ぶ時間を合わせても1200時間程度と推算しており、それが到底十分ではないことを指摘している。つまり、EFL学習者は、意識的に学習時間を作らなければ、十分な英語使用の機会や時間を増やすことはできないのである。この自発的に学習時間を確保して学習する行為が自主学習である。

これまでに英語の自主学習を調べている研究は少なく、大学生の英語自主学習における自己調整学習能力を調査した研究はほとんどない。多くは授業内の個別学習における自己調整学習の効果に着目している。例えば、Chang(2005)は、オンライン授業を通して、学習者の動機づけの変化を自己調整学習の枠組みで調査している。自己調整学習方略を用いたオンライン授業により、学習者の責任感や内発的な動機づけ、難易度の高い問題に取り組む意欲が高くなかったことを示した。また、Liu, Lan, & Ho(2014)は、中国人大学生の英語学習において、オンライン教材による語彙学習は、語彙知識の獲得だけでなく、自己調整学習を促進することが可能であることを結論づけた。Kondo, Ishikawa, Smith, Sakamoto, Shimomura, & Wada(2012)では、任天堂DSを用いたTOEIC対策のモバイル教材の自主学習効果を検証しており、統制群の(モバイル教材が与えられない)大学生よりも、実験群の(モバイル教材を使用した)大学生の方が、有意に自主学習時間が長く、自己調整学習能力が高まるといった結果が得られている。Chang(2005)やLiu et

al.(2014)は、オンライン講義の有効性を検証することが第一の目的であり、教師の介入や指導によって、自己調整学習能力が高くなるということを示唆している。Kondo et al.(2012)は、教師の介入がなくても、大学生英語学習者は主体的に自己調整学習能力を身につける可能性があることを示しているが、モバイル教材はモジュールに従って授業内で使用させている。つまり、いずれの研究も、自主学習に着目している一方で、大学の講義の一部であったり、教育的介入があつたりといった点で、成績や単位と直結した形での英語学習が前提となっている。本研究は、教師の指示や授業の単位などとは関係なく実施する強制力のない自主学習環境に着目している。大学生にとって、真の自己調整学習能力が求められるのは、そういう自発的な学習にあると考えられるためである。最終的に英語を身につけるためには、まず自己調整学習能力を高め、効果的な自主学習を自ら行うことが重要なのである(廣森, 2015; 新多, 2019; 竹内, 2007)。

2.4 自己調整学習能力の調査手法の問題点と研究目的

学習者の自己調整学習能力を理解する研究手法は多岐にわたる。現在は、1980年代に隆盛を極めた質問紙研究に対する個人差要因の軽視といった批判が増え、次第に質的に自己調整学習の変化や過程を捉える研究手法にシフトしている(Schunk & Greene, 2018)。しかし、Wolters & Won(2018)はバラエティに富むデータ収集方法を提案した上で、研究手法の適切性は、研究課題にふさわしいアプローチを取ることで決まるため、自己調整学習の過程(process)を調査するのか、能力(capacity)を調査するのかを入念に考慮すべきとしている。「どのように」自己調整学習を行うのかといった問い合わせた場合、調査の主眼は学習状況に置かれる。つまり、質的研究や混合研究による、タスクや個人差要因、文脈を最大限配慮した研究手法になる。一方、ある集団について「どれくらい」自己調整学習能力があるかといった問い合わせた場合、質問紙調査の意義が見えてくると考えられる。質問紙調査は、データ収集条件の統一が可能であり、多数の協力者に対し一度に実施でき、各協力者の心理的負担が他のデータ

収集方法と比較して小さいことなどから、実施の容易さに関わる利点が極めて大きいと言われている(南風原, 市川, 下山, 2003)。

また、自己調整学習は文脈依存性の高い概念である。これについてWolters & Won(2018)は、自己調整学習研究における尺度の部分的な引用・使いまわしに否定的な立場をとっている。概念自体が学習内容・学習環境・学習中の学習者的情意や行動などに多く依存することから、仮に先行研究の質問紙の一部ないし全部を使用した場合には、統計的処理をもって、客観的な信頼性・妥当性をあらためて確認する必要性を論じている。

以上のことから、本研究では、日本人大学生英語学習者という集団の特性を明らかにすることを目的とし、英語資格試験のための自主学習に特化した自己調整学習能力の尺度開発を試みる。教室環境下を離れた自主学習には様々な目的が伴うが、大学生が授業の予復習以外で自発的に英語学習を行う際、英語資格試験は最も認知されている目的の一つである。この英語資格試験に向けた自主学習環境は、これまで調査されてこなかった領域であり、かつ、そこでどのような自己調整学習が用いられるのかを定量的に明らかにした研究は筆者の知るところない。英語資格試験に向けた大学生の自主学習における「文脈独自の方略」がどのようなものかを、尺度の開発をもって明らかにしたい。

3

質問項目の生成

3.1 材料

質問紙の予備的開発にあたり、次の5つの先行研究をもとに項目プールを作成した。第二言語習得分野だけでなく、教育心理学分野でも用いられてきた複数の質問紙を統合して検討することで、項目の偏りを是正することを意図している。本研究では、自己効力感といった情意要因を重視したZimmerman(1989)を中心とする自己調整学習モデルを理論的背景とした。それぞれの特徴と選考要因を概観する。

3.1.1 Self-Regulated Foreign Language Learning Strategy Questionnaire (SRFLLSQ)

この質問紙はHabok & Magyar(2018)による中学生の教室内外国語学習に特化した自己調整学習能力尺度である。この尺度は、自己調整学習の重要な下位要素である「メタ認知的活動」に焦点を当てたことに独自性がある。彼らは複数の分析過程を経て、「メタ認知」方略の他に「メタ社交(meta-sociocultural-interactive)」方略や「メタ情意(meta-affective)」方略といった、英語圏の文化や対人コミュニケーションを利用した方略や英語学習中に自身の感情を客観的に把握しコントロールする方略も含めて尺度としている。自主学習においてもこれらの方略が用いられる可能性があると判断した。

3.1.2 The Questionnaire of English Self-Regulated Learning Strategies (QESRLS)

Wang et al.(2013)によって作成されたこの尺度は、学習方略に特化した質問紙である。11因子65項目が想定されているが、Wang & Bai(2017)が因子構造を改めて確認したところ1因子解が得られ、妥当性の確認が完全になされたとは言えない。しかし、自己調整学習理論に基づいて項目生成が行われている点、そして、英語学習に特化した学習方略が複数含まれている点(例えば、「発音が似ている単語と一緒に覚える」)から本研究の文脈に合うと判断した。

3.1.3 The Motivated Strategies for Learning Questionnaire(MSLQ)

教育心理学分野・第二言語習得分野を問わず、自己調整学習能力に関する研究において、今日まで最も使用されているのが、MSLQ (Pintrich, Garcia, Smith, & McKeake, 1991, 1993)と呼ばれる尺度である。これは、動機づけと学習方略の2つのセクションで構成されている。動機づけセクションには5つの動機づけ要因、学習方略セクションには5つの認知方略要因と4つの学習資源活用方略が含まれている。自己調整学習能力における動機づけ的側面の重要性を説いた本質問紙の意義は大きく、動機づけの基盤があることで

効果的な自己調整学習方略が使用されるという立場にある。

3.1.4 Metacognitive Awareness

Inventory (MAI)

これは、メタ認知の理論に基づいた尺度であり、Schraw & Dennison(1994)によって提案されている。MAIは外国语学習には特化しておらず、内容理解度の把握や学習の管理といった一般的な学習を想定した項目で構成されている。自己調整学習方略の大部分はメタ認知方略によって支えられていることから参考とした。

3.1.5 動機づけ調整方略尺度

梅本・田中(2012)によるこの尺度は、「やる気が出ないときにどのようにしてやる気を起こすか」という状況に特化して作成されている。これまでの自己調整学習に関する質問紙では行き届かなかった、学習者が自身の動機をどのように調整するかといった観点を持つ質問紙だと言える。本研究の対象である英語資格試験に向けた自主学習という学習環境においては、減退したやる気を自分自身でコントロールする方略の必要性が考えられるため参照した。

3.2 手続き

まず、上記の質問紙に含まれる項目を「英語資格試験に向けた自主学習」に相応する内容および表現に改訂した。英語資格試験は実施団体によってさまざまな形態で提供されているが、調査対象者(日本人大学生)にとって特に馴染みがあると考えられる「英検」「TOEIC」「TOEFL」「IELTS」の4試験を「英語資格試験」と定義した。次に、大学生の時に英語資格試験を受験したことのある大学院生4名にインタビューを実施し、自主学習経験および項目プールの内容に関するデータを収集した。その後、自己調整学習の理論的背景と照らし合わせ、6つの測定される要素を仮定した。
①メタ認知方略、②学習行動管理方略、③メタ情意方略、④自己効力感、⑤目標志向性、⑥動機づけ調整方略である。これらを前提に計300項目の統合・除外を行った。作業は博士課程の大学院生2名と調査者の計3名で繰り返し行われ、最終的に

動機方略として26項目、動機づけ調整方略として15項目、学習方略として81項目の計122項目が選ばれた。

4 質問紙調査の実施

4.1 協力者

都内の大学に通う大学1年生から4年生356名(男性97名、女性257名、無回答2名)が調査に参加了。彼らの専攻は、福祉学(82名)、現代心理学(76名)、社会学(58名)、経済学(40名)、理学(37名)、人文科学(31名)、法学(19名)、経営学(6名)、異文化コミュニケーション学(6名)、観光学(1名)である。大学生になってから1度以上英語資格試験(英検・TOEIC・TOEFL・IELTS)を受験したことがある大学生に協力を募り、TOEICの受験経験があることを条件とした。調査段階においては、英検の受験経験者が56名(15.7%)、TOEFLの受験経験者が6名(1.7%)、IELTSの受験経験者が18名(5.1%)であった。また、英語圏への留学経験については、ほとんどの大学生(303名)はない回答したが、20名が1か月未満、21名が1か月以上3か月未満、7名が3か月以上1年未満、5名が1年から4年の留学を経験している。

4.2 手続き

質問項目はウェブアンケート「Questant」を用いて作成された。大学内で募集を行い、参加意思を示した学生のメールアドレスにアンケートのURLを送付し、回答を得た。データ収集は2020年1月15日から1月31日の2週間で行われ、協力者に対する研究協力への同意確認は、募集時、URLの送付時、ウェブアンケートの開始時の3回にわたって行われた。順序効果を避けるためアンケートの質問項目はランダムに表示され、協力者は5件法(1:あてはまらない、2:あまりあてはまらない、3:どちらともいえない、4:ややあてはまる、5:あてはまる)によって、最も自身を表現している数値を選ぶものとした。教示文は、「英語資格試験に向けて自主学習をするとき、あなたの考え方や行

動について教えてください」とし、動機づけ調整方略尺度に関する項目については、「あなたは英語資格試験に向けた自主学習を行う中でやる気が出ないとき、どのようにやる気を起こしていますか?」とした。いずれも、大学生になってからの英語資格試験に向けた自主学習を想定して回答するよう強調した。

4.3 分析

自己調整学習能力を3つの尺度からなる概念として、3.2手続きにある6つの要素を、①から③は学習方略尺度、④と⑤を情意尺度、⑥を動機づけ調整方略尺度と位置づけ、個別に分析することとした(注1)。分析は、統計ソフトSPSS(ver. 26.0)およびAmos(ver. 26.0)を使用した。分析実施の前提として各質問項目に問題がないかを判断するための記述統計、各質問項目の共通性を探るための探索的因子分析、探索的因子分析の結果が構造的に妥当かを検討するための確証的因子分析、得られた因子と自主学習頻度および資格試験得点との関係性を検討するための相関分析と回帰分析、そして、得られた因子間の関係性を評価する共分散構造分析を用いた。特に、探索的因子分析ではより客観的な方法で因子数を決定するために、統計フリーソフトR(ver. 3.5.1)を用いてMAP(Minimum average partial)基準、BIC(Bayesian information criterion)指標、平行分析を併用し、信頼性係数の計算では α 係数の他に ω 係数を算出した。

5 結果

5.1 探索的因子分析による質問項目の精査

5.1.1 情意尺度の因子構造と命名

自己調整学習の情意的側面を測定するための26項目に対し、探索的因子分析を行った。MAP基準では2因子、BIC指標は3因子、平行分析では4因子が示唆されたが、固有値の減衰状況を考慮し3因子解が妥当と判断した。因子負荷量の最低値を.40とした上で、探索的因子分析(最尤法・プロマックス回転)を実施し、表1の推定結果が得られた。

第1因子は「自分なら資格試験で理想の点数が取れると思う」「試験を実際に受けるときは、自分の力を出し切れる自信がある」といった、自分に対する期待や自信を示す5項目によって構成されていることから「自己効力感」と命名した。第2因子は、「勉強内容が十分にわかったら、とても満足すると思う」「英語を勉強することは役に立つと思う」といった、英語の習得そのものに価値や目的を示す項目が並んでいることから「内発的目標志向」と命名した。対して、第3因子は、「良い結果を残して、自分の実力を周りの人々に自慢したい気持ちがある」といった、試験結果の良さや、資格試験の対外的な成功を示す項目で構成されていることから「外発的目標志向」と命名した。これら3因子の因子寄与率は33.37%であった。

5.1.2 動機づけ調整方略尺度の因子構造と命名

動機づけ調整尺度は、15項目のうち4項目が項目分析によって除外され、11項目で実施された。推定する因子数を分析したところ、MAP分析では1因子、BIC指標と平行分析では2因子が示唆された。固有値の減衰状況もあわせて考慮した結果、2因子解を採用することとした。そこで、各因子に対する因子負荷量を.40に設定し、それより負荷量の低い項目は削除しながら探索的因子分析(最尤法・プロマックス回転)を行った結果、表2の通りの因子構造が得られた。第1因子には、「早いうちに一定のスコアを取ってしまったほうが楽」「就職に有利」などの5項目が含まれており、資格試験に向けた自主学習へのやる気が出ないときに、考え方を変えることでやる気を起こす方略であることから「認知調整」と命名した。第2因子は、「勉強の合間に気分転換をして、やる気を保つ」といった学習環境に着目した動機づけの調整行動を示す方略で構成されていることから「環境調整」と命名した。これら2因子の寄与率は28.83%であった。

■表1: 情意尺度の探索的因子分析結果

番号	項目	自己効力感	内発的目標志向	外発的目標志向
M013	自分なら、理想の点数(級)の取得(合格)ができると思う。	0.855	-0.095	-0.025
M017	英語の資格試験の受験は、きっとうまくいくと思っている。	0.659	-0.035	0.021
M014	今のところは到底不可能な目標(点数や級)もいずれ到達できると思う。	0.578	0.104	-0.067
M015	英語の資格試験に向けて取り組む練習問題が難しくても乗り越える自信がある。	0.476	0.104	0.003
M016	試験を実際に受けるときは、自分の力を出し切れる自信がある。	0.410	0.013	0.009
M03	勉強内容が十分にわかったら、とても満足すると思う。	0.013	0.649	0.010
M011	もし点数が悪かったり、不合格だった場合は、自分のせいだと思っている。	-0.090	0.517	0.002
M07	英語の資格試験に向けて勉強した内容や経験そのものが、英語力の向上につながると思う。	0.059	0.505	0.010
M09	英語を勉強することは役に立つと思う。	0.126	0.462	-0.012
M06	英語資格試験で良い結果を残して、自分の実力を周りの人に自慢したい気持ちがある。	-0.064	-0.030	0.638
M04	英語資格試験の結果がとにかく大事なので、そのため勉強する。	-0.126	0.104	0.499
M026	英語の資格試験を受けることで、周囲の期待に応えていると感じている	0.082	-0.088	0.477
M05	他の人より英語で良い点数や良い結果を残したいので自主学習をする。	0.225	0.032	0.467
	因子間相関	0.215	0.424	
				0.135

■表2: 動機づけ調整方略尺度の探索的因子分析結果

番号	項目	認知調整	環境調整
MR5	早いうちに一定のスコアを取ってしまったほうが楽だと考える。	0.689	-0.195
MR15	資格試験のための勉強は自分に必要なことだと言い聞かせる。	0.584	-0.051
MR10	資格試験の受験が就職に有利だと考える。	0.491	0.104
MR4	英語の成績を良くするためだと考える。	0.440	0.117
MR2	資格試験の受験に役立ちそうな情報を探す。	0.400	0.221
MR13	勉強の合間に気分転換をして、やる気を保つ。	0.033	0.521
MR12	勉強場所を変える。	-0.108	0.506
MR1	計画を変更して、少しでも良いから進めようとする。	0.068	0.469
	因子間相関	0.391	

5.1.3 学習方略尺度の因子構造と命名

学習方略を測定するための81項目のうち、「単語は語呂合わせで覚える」など4項目に床効果(注2)や過度のばらつきが見られたため、それらを除外した77項目に対し探索的因子分析を行った。MAP基準は5因子、BIC指標は3因子、平行分析は7因子を示唆したが、3因子解と7因子解では不適解が生じたため、固有値の減衰状況を考慮した上で5因子解が妥当と判断し、.40を因子負荷量の最低値として、探索的因子分析(最尤法・プロマックス回転)を行った。因子の推定結果を表3に示す。第1因子は、「やらなければいけない場面では、自分自身をやる気にさせることができる」

「資格試験に向けた勉強がつまらなくなったり退屈になったりしても、自分で決めたところが終わるまでは作業を続ける。」といった、自分自身を律して試験勉強に取り組むといった方略で構成されているため「努力方略」と命名した。第2因子は、「勉強中にわからないことがあったら、自分の知識だけで解決しようとせず、教材・辞書・インターネットなどで調べる」「知らない単語が出てきたとき、それがわからなくとも内容理解に支障がない場合は、飛ばして読む」からわかるように、難しい内容やわからない問題が出てきた場合の対応や適応に関する項目を含むことから「対処方略」とした。第3因子は、「発音が似て

いる英単語と一緒に覚える」や「文法を整理するために表を作る」といった、語彙や文法などの暗記項目の覚え方に関する5項目で構成されているため「記憶方略」と命名した。第4因子は、「自分から進んで英語の動画を見ようとする」「進んで英語のニュースや音楽などの音声を聞くようにしている」といった、資格試験前に関わらず、日頃から継続的に英語に触れるための方略である。

ることから「日常的方略」と命名した。第5因子は、「自分が受ける資格試験の模擬問題を、時間を測りながら解く」や「英語の学習だけでなく、自分が受けるテスト形式の確認も十分にする」など、資格試験に特化した対策や計画を含む5項目で構成されているため「試験計画方略」と命名した。なお、因子寄与率は37.13%であった。

■表3: 学習方略尺度の探索的因子分析結果

番号	項目	努力 方略	対処 方略	記憶 方略	日常的 方略	試験計画 方略
SR73	やらなければいけない場面では、自分自身をやる気にさせることができる。	0.715	0.046	-0.274	-0.110	0.043
SR6	勉強中は自分自身を奮い立たせてやるべきことをやっている。	0.636	0.057	-0.042	-0.076	0.029
SR21	資格試験に向けて勉強しているときは集中力が続くほうだ。	0.590	0.002	0.061	0.056	-0.045
SR31	資格試験に向けた勉強がつまらなくなったり退屈になったりしても、自分で決めたところが終わるまでは作業を続ける。	0.580	0.146	0.023	-0.108	-0.076
SR73	試験勉強は、自分で決めたことだから楽しんでできていると思う。	0.510	0.014	0.138	0.171	-0.180
SR29	計画的に勉強しているので、試験直前に焦ることはほとんどない。	0.501	-0.224	0.050	0.024	-0.039
SR28	他の活動が忙しい中でも、勉強時間を作って自主学習をするよう心掛けている。	0.496	0.078	0.080	0.112	0.052
SR27	自分自身に何らかのノルマを課して勉強している。	0.424	-0.099	0.089	0.008	0.250
SR47	勉強中にわからないことがあったら、自分の知識だけで解決しようとせず、教材・辞書・インターネットなどで調べる。	-0.041	0.643	-0.196	0.086	0.017
SR62	知らない単語の意味を、文脈から推測する。	0.082	0.620	-0.094	-0.028	0.007
SR63	知らない単語が出てきたとき、それがわからなくとも内容理解に支障がない場合は、飛ばして読む。	0.032	0.502	-0.163	0.030	0.095
SR39	英語を聞くときは、内容を理解するために強調されている語やフレーズに注意を向ける。	0.027	0.498	0.237	0.044	0.001
SR49	1回で聞き取れなかったら、何回も英語の音声を聞き返す。	0.040	0.463	0.164	-0.020	-0.023
SR48	1回読んで内容が理解できなかったら、何回も英文を読む。	-0.037	0.408	0.161	0.039	0.037
SR66	以前うまくいった勉強方法で勉強する。	0.004	0.404	0.077	-0.145	0.088
SR44	発音が似ている英単語と一緒に覚える。	-0.073	0.151	0.639	0.019	-0.032
SR43	単語の細かい意味の違いを区別するために、類義語も一緒にまとめて覚える。	-0.049	0.165	0.611	-0.053	0.062
SR38	覚えやすいように単語を分類する。	0.010	-0.154	0.583	-0.002	0.094
SR42	文法を整理するために表を作る。	-0.004	-0.196	0.485	-0.074	0.043
SR59	英語と日本語の類似点や相違点を比べる(例:文法・意味・単語)。	0.081	0.074	0.472	-0.029	-0.145
SR50	自分から進んで英語の動画を見ようとする。	-0.061	0.009	-0.071	0.813	-0.019
SR51	進んで英語のニュースや音楽などの音声を聞くようにしている。	-0.016	-0.037	-0.107	0.769	0.066
SR2	出来る限り英語で書かれたものを読む機会を設ける。	0.030	0.038	0.103	0.670	-0.045
SR80	自分が受ける資格試験の模擬問題を時間を測りながら解く。	-0.143	0.098	0.017	0.029	0.663
SR78	英語の学習だけでなく、自分が受けるテスト形式の確認も十分にする。	-0.003	0.105	-0.073	0.018	0.598
SR79	テストを受けるときの時間配分を考える。	-0.008	0.047	-0.018	-0.091	0.593
SR55	英語の資格試験に向けた勉強計画を立てる。	0.205	-0.136	0.072	0.111	0.466
SR56	必ず目標(例:○○点を取る・○級に合格する)を設定する。	0.153	0.068	0.068	0.021	0.445
	因子間相関	0.072	0.403	0.408	0.429	
			0.090	0.035	0.436	
				0.391	0.228	
					0.226	

5.1.4 下位尺度得点間の相関係数と記述統計量

情意尺度、動機づけ調整方略尺度、学習方略尺度の下位尺度得点における相関係数と、平均値、標準偏差、信頼性係数(α , ω)を表4に示す。従来の自己調整学習研究では、情意要因と学習方略には正の相関があり、また、動機づけ調整方略

と学習方略にも正の相関があることが認められてきた。本研究では全体的に、情意尺度、動機づけ調整尺度、学習方略尺度が互いに正の相間にあり、認知調整、試験計画方略の2因子については、すべての因子と有意な正の相関関係にあることがわかった。

■表4: 下位尺度得点間の相関係数と記述統計量

	情意尺度			動機づけ調整方略尺度		学習方略尺度								
	自己効力感	内発的目標志向	外発的目標志向	認知調整	環境調整	努力方略	対処方略	記憶方略	日常的方略	試験計画方略	M	SD	α	ω
自己効力感		0.191***	0.309***	0.203***	0.079	0.451***	0.199***	0.184***	0.404***	0.326***	3.29	0.786	0.73	0.74
内発的目標志向			0.105*	0.305***	0.303***	0.093	0.517***	0.018	0.142**	0.284***	4.36	0.577	0.61	0.62
外発的目標志向				0.302***	0.104	0.342***	0.152**	0.205***	0.152**	0.333***	3.22	0.881	0.61	0.61
認知調整					0.244***	0.173**	0.240***	0.194***	0.122*	0.318***	3.79	0.802	0.66	0.67
環境調整						0.146**	0.324***	0.127*	0.071	0.206***	3.98	0.802	0.47	0.50
努力方略							0.126*	0.302***	0.318***	0.397***	3.04	0.764	0.79	0.79
対処方略								0.140**	0.066	0.403***	4.21	0.558	0.72	0.73
記憶方略									0.231***	0.199***	2.73	0.875	0.68	0.68
日常的方略										0.209***	2.99	1.130	0.78	0.78
試験計画方略											3.80	0.791	0.72	0.73

(注)* $p < .05$, ** $p < .01$, *** $p < .001$

5.2 信頼性・妥当性の検証

次に、5.1で得られた3尺度に十分な信頼性および妥当性があるかを検証した。Messick(1995)は6つの評価基準を示しているが、必要に応じて複数の評価基準を採用することが推奨されていることから、本研究では、①それぞれの因子に含まれる項目が同じ概念を測定しているか、②尺度の因子構造が客観的に適切と判断できるか、③作成した尺度は外的基準と相関関係や因果関係にあるか、の3点を検討する。

5.2.1 内的整合性

内的整合性とは、情意尺度、動機づけ調整方略尺度、学習方略尺度の各因子に含まれている項目の類似性を指す。本研究では、クロンバッックの α とマクドナルドの ω (注3)を指標とし、因子のまとまりを検証した(表4)。情意尺度の3因子では、 $\alpha = .61 \sim .73$, $\omega = .61 \sim .74$ 、学習方略尺度の5因子では、 α 係数および ω 係数ともに $.68 \sim .79$ が得られており、良好な内的一整合性が得られた。一

方、動機づけ調整方略尺度については、認知調整の内的信頼性は保たれているものの($\alpha = .66$, $\omega = .67$)、環境調整は、 $\alpha = .47$, $\omega = .50$ と不十分な数値であった。しかし、動機づけ調整方略は特定の学習状況に依存しやすく、先行研究でも同様の傾向が確認されている。本研究でも、環境調整因子に含まれる3項目は、理論的背景を踏まえた上で必要な項目として生成されていることから、このまま採用することとした(6.2参照)。

5.2.2 構造的側面の妥当性

情意尺度、動機づけ調整方略尺度、学習方略尺度のそれぞれの因子構造は、探索的因子分析によっていわば研究者の主観的な判断に依拠して得られたものである。そこで、得られた因子構造が客観的な評価からも適切と判断できるかを調べるために、確証的因子分析を実施した。その結果を表5に記載している。情意尺度、動機づけ調整方略尺度、学習方略尺度のいずれも十分な適合度が得られた。具体的には、①モデルがどの程度うまく説明できているかを示す指標(GFI・

AGFI), ②モデルがどの程度改善されたかを評価する指標(CFI), ③モデルがいかに悪いかを算出する適合度指標(RMSEA)から判断した(注4)。つまり、探索的因子分析によって得られた因子構造は、理論的仮説を照らし合わせても相反しな

いものであり、確証的因子分析の結果からも尺度構造の妥当性が担保されたと言える。

■表5: 3つの尺度の確証的因子分析結果

	CMIN	df	p	GFI	AGFI	CFI	RMSEA
情意尺度	66.354	56	0.162	0.972	0.955	0.985	0.023
動機づけ調整方略尺度	17.366	16	0.362	0.988	0.973	0.996	0.016
学習方略尺度	336.973	301	0.075	0.935	0.913	0.984	0.018

(注) CMIN = カイ2乗値, df= 自由度, p= 有意水準, GFI・AGFI・RMSEA = 注4を参照

5.2.3 外的側面の妥当性

当該3尺度と外的基準との間に相関があるかを測るために、相関分析を実施した。仮説では、自己調整学習と関連のある変数は正の相関関係にあることが想定されている。外的基準として本研究では、現時点での資格試験得点(TOEICのスコア)と、資格試験に向けた自主学習の頻度(全くしない・あまりしない・たまにする・必ずする)に関するデータを得ており、いずれも自己調整学習能力とは正の相関関係が予想される。

分析の結果(表6)、資格試験得点はいわば協力者の調査時点での英語力としてみなすことができるが、負の相関関係にある尺度因子はなく、仮説をおおよそ支持できる。ただし、有意な相関関係にあったのは、自己効力感、外発的目標志向、対処方略、日常的方略の4因子のみであった。つまり、資格試験得点の高さは、自分に自信があることや、試験結

果を重視する目標志向性、わからない問題に遭遇したときにうまく対処できること、そして、日常的に英語を読んだり、聞いたりする習慣と関係があるということがわかった。一方、その他の因子は、ほとんど無相関に近いものが多く、資格試験得点との関連が強くないことを示す結果となった。

自主学習の頻度についても、負の相関関係は見られず、記憶方略、日常的方略以外のすべての因子と有意な相関関係があった。記憶方略は、暗記や認知作業に特化した方略であり、自主学習頻度との関連が想定されない特性があり、日常的方略に關しても、自主学習としてではなく趣味や習慣として行われる方略であることも鑑みると、有意な相関関係はないことが理解できる。そのため、本研究の尺度は、自己調整学習態度があるほど、自主学習の頻度が高いという仮説を支持する相関関係が示された。

■表6: 下位尺度得点と外的基準の相関係数

	自己効力感	内発的目標志向	外発的目標志向	認知調整	環境調整	努力方略	対処方略	記憶方略	日常的方略	試験計画方略
資格試験得点	0.238***	0.089	0.116*	0.041	0.034	0.058	0.148**	0.017	0.329***	0.079
自主学習頻度	0.125*	0.227***	0.207***	0.221***	0.212***	0.232***	0.288***	0.018	0.080	0.438***

(注)*p < .05, ** p < .01, *** p < .001

続いて、得られた因子が外的基準をどの程度予測するのかを、重回帰分析を用いて検証した。自主学習頻度を目的変数とした場合と、資格試験得点を目的変数とした場合の重回帰分析の結果を以下にまとめると。

情意尺度については、調整済みR²を見ると、自主学習の頻度の7.9%、資格試験得点の5.3%を有

意に予測することがわかった。表7をみると、英語学習に対する内発的・外発的目標志向を強く持っているか否かが、自主学習の頻度に影響を与えることが示された。一方、資格試験の得点に対しては、自己効力感の高さが影響を与えるということが示された。

表7: 情意尺度の重回帰分析結果

	自主学習頻度			資格試験得点		
	B	CMIN	β	B	CMIN	β
自己効力感	0.034	0.058	0.032	23,431	7,232	0.217**
内発的目標志向	0.295	0.076	0.202***	7,602	9,426	0.042
外発的目標志向	0.168	0.051	0.176*	5,215	6,370	0.045
調整済み R^2	0,079***			0,053***		
N	356			356		

(注)B = 推定値, SEB = 標準誤差, β = 標準回帰係数, * $p < .05$, ** $p < .01$, *** $p < .001$

動機づけ調整方略尺度に関しては、自主学習頻度にのみ有意な説明力が見られ、7%の説明率であることがわかった(表8)。やる気が起きないときに、考え方を変えたり環境を変えたりすることが自主学習頻度を増やすと考えられる。一方、資格試験の得点に対しては、有意な因果関係が見られなかった。つまり、自身のモチベーションややる気を維持するために認知や環境を調整する方

略は、資格試験得点の高低には直接的な影響は及ぼさないことがわかった。しかし、動機づけ調整方略は、学習過程における刻一刻と変化する動機づけを安定させ、積極的な学習行動を促す点で効果的であり(梅本・稻垣, 2019), その学習の成果が資格試験得点に現れると考えると、間接的な影響力がある可能性が考えられる。

表8: 動機づけ調整方略尺度の重回帰分析結果

	自主学習頻度			資格試験得点		
	B	SEB	β	B	SEB	β
認知調整	0,189	0,055	0,18**	4,436	7,054	0,034
環境調整	0,177	0,055	0,169**	3,319	7,055	0,026
調整済み R^2	0,070***			-0,003		
N	356			356		

(注)* $p < .05$, ** $p < .01$, *** $p < .001$

学習方略尺度の影響力を重回帰分析で検証したところ、自主学習頻度に対して21%, 資格試験得点に対して12.1%の影響力をもつことがわかった(表9)。自主学習頻度に対しては、対処方略、記

憶方略、試験計画方略が有意である一方、資格試験得点に対しては、対処方略と日常的方略が有意であることが示された。自主学習の頻度と資格試験の得点に影響を及ぼす要因はそれぞれ異なっ

表9: 学習方略尺度の重回帰分析結果

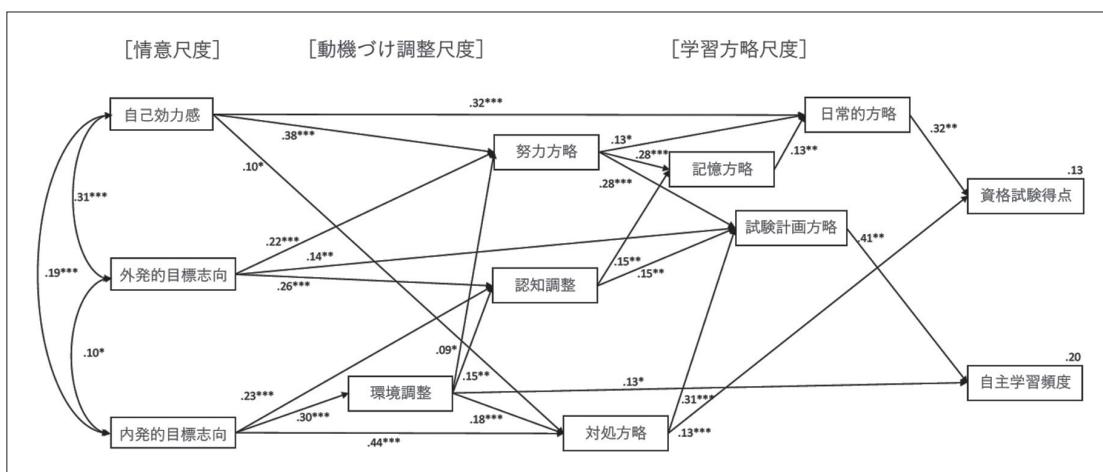
	自主学習頻度			資格試験得点		
	B	SEB	β	B	SEB	β
努力方略	0,117	0,060	0,106	-5,731	7,770	-0,042
対処方略	0,218	0,078	0,144**	27,700	10,100	0,150*
記憶方略	-0,099	0,048	-0,103*	-8,031	6,266	-0,068
日常的方略	-0,011	0,038	-0,015	32,317	4,867	0,354***
試験計画方略	0,385	0,060	0,361***	-3,289	7,721	-0,025
調整済み R^2	0,210***			0,121***		
N	356			356		

(注)* $p < .05$, ** $p < .01$, *** $p < .001$

ていることと、対処方略はどちらに対しても影響力を持つのに対し、努力方略はどちらにも有意な影響がないということが示された。以上の結果から、本研究で開発された尺度は、外的基準である自主学習頻度と資格試験得点との間に一定の相関関係と因果関係があることが示され、構成概念妥当性が確認されたと言える。

最後に、すべての因子と資格試験得点および自主学習頻度との関係性を調べるためにパス解析を行った。分析に先立ち、3つの尺度の因果関係について仮説を立てた。先行研究では、動機づけや情意といった意欲に関わる因子は、学習方略使用の前提として不可欠であるとされている(例えば、Kormos & Csizer, 2014; Schunk &

Greene, 2018; Ushioda, 2014)。動機づけ調整方略尺度の2因子は、「方略」であることから学習方略と同様の扱いとし、相関分析や重回帰分析の結果をもとに因果関係を予測した。その結果を、図1にまとめている。仮説通り、情意要因は方略に先立つ要因として考えられ、自己効力感は学習方略に対し、2つの目標志向性は学習方略だけでなく動機づけ調整方略に対しても有意なパスが見られた。資格試験得点は日常的方略と対処方略からのみ、自主学習頻度は試験計画方略と環境調整方略からのみ、有意なパスを得るという結果になった。さらに、自主学習の頻度と資格試験得点の間には因果関係が見られなかった。



■図1:パス解析の結果

(注)CMIN = 54.567, df = 40, p = .062, GFI = .975, AGFI = .952, CFI = .981, RMSEA = .032, SRMR = .038, AIC = 130.567.
* p < .05, ** p < .01, *** p < .001, 誤差項は省略した。

6

解釈と議論

本研究は、自己調整学習能力尺度を作成し、その尺度の信頼性および妥当性を検証することが目的であった。複数の統計的処理を経て、3つの尺度の安定性と一貫性は認められたと言える。

6.1 情意尺度の因子構造と外的基準との関連

探索的因子分析の結果、情意尺度には自己効力感、内発的目標志向、外発的目標志向が含まれ

ることがわかった。先行研究では、十分な自己効力感や内発的目標志向が自己調整学習に正の影響を及ぼすことが認められており、とりわけ自己効力感の存在は、自己調整学習が認知的成果(学業成績など)、情意的成果(楽しさなど)、そして動機づけ的成果(動機の継続など)につながる前提条件として、なくてはならないものとして理解されてきた(Lens & Vansteenkiste, 2008)。一方、外発的目標志向、すなわち、他者と比較して自分が優れていることへの価値や、試験結果への強いこだわりといった要素は、内発的目標志向と比較して、学習者が失敗への恐怖感を強めたり、他者への援助要請をしなくなるといった負の影

響を与えるものとして考えられてきた(Fryer & Elliot, 2008)。しかし、本研究のような英語資格試験を目的とした文脈において、この外発的目標志向が学習者に強く内在していることは、特筆すべき知見といえる。自己調整学習理論においては、学習の原動力が自らの知的好奇心や達成感などを満足させる(内発的な)ものにあるほうが、他者に対する優位性や外的評価の高さを満たす(外発的な)ものにあるよりも効果的であると考えられてきた。そのため、外発的目標志向を必須としている理論は少ない。しかし、英語資格試験を念頭に置いた自主学習という性質上、外発的な価値が重視されるのは当然のことと考えられる。日本人大学生にとって英語資格試験とは、常に受験できるものではなく、また、それによって将来的のライフィベント(留学や就職など)に影響を与える試験である。つまり、受験料や参考書の元を取るといった損得勘定が働いたり、資格として履歴書に書かなければならないといった過剰不安が働いたりする可能性がある。そのため、外発的目標志向が自身の学習を管理・コントロールするという点では、本研究の結果は必然と考えられる。例えば、韓国の大学生を対象に彼らの理想自己(Ideal L2 Self)を調査したKim & Kim(2014)が示唆しているように、英語力の証明が社会的な評価を左右しうる影響力をもつ韓国では、義務自己(Ought-to L2 self)と呼ばれる外発的な動機づけが学習者に強く作用することがわかっている。彼らを参考にすると、外発的目標志向が自己調整学習能力の一要因として抽出された点を解釈する上で、英語資格試験の受験にあたり学習者には「～しなければならない(義務自己)」といった動機づけが働いていると推察される。

6.2 動機づけ調整方略尺度の因子構造と外的基準との関連

動機づけ調整方略尺度では、やる気が起きないときの調整方略として、自分に有益なことを考える認知調整と、学習環境を変える環境調整の2因子が抽出された。いずれの信頼性(α 係数・ ω 係数)も低いことが本研究の課題である。しかし、梅本・田中(2012)でも、本研究と類似の項目で構成された「認知変容方略」「環境調整方略」因子が抽出されており、低めの α 係数が報告されている

(それぞれ順に、 $\alpha = .57$, $\alpha = .59$)。これは、帯域幅と忠実度のジレンマと呼ばれる、信頼性を高めすぎることで項目同士が類似しすぎる一方、妥当性を高めるために幅広い内容の項目で構成することで項目の共通性が低くなりがちになってしまう現象が起きていると考えられる。梅本・田中(2012)はこれを学習方略の状況依存性によるものとして解釈している。つまり、学習状況に応じて項目を生成すると、扱いきれないほどの項目数を抱えることになるため、必要最低限の項目数に減らさざるを得ず、それが信頼性の低さを招いたと考えられる。この点で、本研究で検討された動機づけ調整尺度は改良の余地があると言える。

ただし、動機づけ調整方略は、情意尺度の内発的目標志向、外発的目標志向との間に有意な因果関係があり、学習方略尺度に対しても有意な因果関係が見られた。さらに、環境調整によってやる気を起こす行為は、自主学習頻度にも影響を与えるということが明らかになった(図1)。このことから、動機づけ調整方略は、資格試験得点との直接的な関係性はほとんど見られなかつたが、自主学習をすることに関しては、重要な役割を果たすと言える。今後は、「やる気が出ないときにどのようにやる気を起こすか」といった方略を用いる英語学習場面について、質的なアプローチで検証することにより個人差を明らかにすることが可能となり、質問紙開発に対する視座を与えるものと思われる。

6.3 学習方略尺度の因子構造と外的基準との関連

自己調整学習における方略的要素をまとめて学習方略とし、努力方略、対処方略、記憶方略、日常的方略、試験計画方略の5因子が抽出された。英語資格試験に向けた学習方略として、これらは非常に特徴的であると言える。まず、努力方略および対処方略は、極めてメタ認知的側面が強い方略である。特に学習中に、自分自身の集中力ややる気を維持し、自分自身が何を理解して何をわかっていないのか、そしてそれをどのように解決するのかといった方法を模索することは、学習者に高度な客観的認知能力を要求する行為である。また、日常的方略は、資格試験得点に最も影響を与える方略として抽出されたが($\beta = .345$, $p <$

.001), 日常に英語に触れようとする姿勢はそれ自体が一つの自己調整学習サイクルとして機能していると言える。英語資格試験の前に自主学習を行う他, 日頃から自身の動機や行動を維持, 管理しているからこそ, 習慣的に英語に触れる環境が構築できていると考えられる。さらに, 記憶方略が抽出され, 資格試験を受験するための具体的な記憶方法があることがわかった。特にこれらは語彙習得と関連しており, 暗記のためのさまざまな工夫が方略として見られる。記憶方略は自主学習頻度や資格試験得点とは直接的な関係を示さなかったが, どの方略を使用するかだけでなく, どのように組み合わせて使用するかが重要といったS²R理論の指摘を反映するものとなつた(Oxford, 2011, 2017)。試験計画方略もメタ認知能力を具現化した方略と言えるが, これは特に, 本研究の調査対象である「資格試験に向けた自主学習」という文脈に特化した因子だと言える。自分が受験する資格試験の受験方法, 試験内容, 試験時間を, 事前に練習したり, 計画的に把握したりするといった能力が, 自主学習をする上で鍵となる方略であるといえる。この試験計画方略は, 自己調整学習理論において, 予見(foreshadowed) (Zimmerman, 2000) という行為に分類され, 自己調整学習の始発段階における方略だと考えられる。予見段階では, 計画立案や目標設定の他に, 情意尺度にある自己効力感や内発的・外発的目標志向といった自己期待や自己信念が機能するとされている。概して, 本研究では, 効果的な自己調整学習のための第一歩として重要な役割を担う要素が測定されたことが確認できた。

6.4 資格試験に向けた自主学習における自己調整学習の役割

パス解析の結果(図1参照), 自主学習と資格試験得点に対する自己調整学習の関わり方にについていくつかの知見が得られた。1つ目は, 自己調整学習方略は, 資格試験得点にも自主学習頻度にも影響を与えるが, 直接的に関与する方略は, それぞれで異なっている点である。資格試験得点には日常的方略と対処方略からのみ, 自主学習頻度には試験計画方略と環境調整方略からのみ有意なパスが認められた。つまり, 資格試験得点であるTOEICのスコアの高低には, 日常的に英語

に触れているかどうかと, わからない問題や知らない知識について, それを解消するためにどう取り組むか, が関わっているということが言える。一方, 資格試験前の自主学習では, 自己調整学習方略の中でも, 試験に向けた計画や試験形態そのもののへの順応といった方略と, やる気が起きないときに環境を調整してやる気を奮い立たせるといった方略が関係することが明らかとなった。すなわち, 学習習慣を確立させるための自己調整学習と試験得点を上昇させるための自己調整学習は, 性質が異なるということが示唆された。

2つ目は, 本研究の参加者においては, 自主学習頻度と資格試験得点に相関関係が見られなかつたことである。これは尺度開発そのものとは別に考慮すべき点であるが, 本研究の結果からは, 資格試験に向けてどれだけ自主学習したとしても資格試験の得点にはほとんど影響がない, という解釈が成立する。ただし, 本研究の尺度は英語資格試験(4つの試験)を前提としているのに対し, 資格試験得点はTOEICのスコアのみが反映されている。参加者が複数の異なる英語資格試験を想定して自主学習頻度を回答した可能性があり, 特定の試験得点との関連が弱かったと考えられる。

3つ目は, 情意尺度因子である自己効力感, 内発的目標志向, 外発的目標志向はすべて, 方略に先立つ前提条件として存在することが確認され, 先行研究を支持する結果となった。すなわち, 学習方略を効果的に活用するためには, あらかじめ自己効力感と適切な目標志向を有することが基盤となる。自己効力感の重要性とともに内発的目標志向の意義は語られてきたが, 資格試験を受験する上では, 外発的目標志向が重要な役割を果たしうる。達成目標理論(Elliot & McGregor, 2001)では, 外発的目標志向はパフォーマンス目標に相当し, 「他者と比較して自分の能力や優秀さをアピールしたい」といった接近型と「他者と比較して自分の出来なさや無能さをさらすことを避けたい」といった回避型に細分化されるとしている。自己調整学習において有効なのは, 前者であろう。自らの能力の高さを証明したいといった積極性を持つ目標志向が本研究の尺度にも反映されている。

これらを総合的に解釈すると, 自分なら上手く

いく・やり遂げられるといった自己期待、英語の習得自体への目標、および、結果や他者との競争に対する目標を持つことは、資格試験に向けた自主学習における原動力となる。それが、さまざまな方略 一例えれば、日常的に読んだり聞いたりして英語に触れること、わからない問題にどのように対処するべきか身につけること、具体的で自分にとって重要な目標を定めること、そして、それを達成するためにはどのような時間配分でどのような方法で試験問題を解くかといった戦略に焦点を当てること一と複雑に結びつき、自主学習や資格試験得点につながっていくと考えられる。資格試験の受験に向けた自己調整学習能力として、これらを体系的に組み合わせて実行に移すことが肝要であると言えるだろう。

本研究の対象は、教室環境下を離れた自主学習としているが、これらの知見は、教育現場でも大いに活用されうると考える。自己調整学習方略の指導効果は多く検証されており、方略の種類や使い方を明示的に指導することは大いに有益だと言える。ただし、これらを活用するためには情意的基盤が必須であり、具体的には、教師は学習者の自己効力感と目標志向性を育み支援することが必要である。特に、これまでのような内発的な目標志向性を培う教育的介入とともに、大局的な見地に立ち、学習者の多様な目標の在り方を尊重し支援する必要性があることを示唆したと言える。竹内(2007)は、教師からの働きかけによって、学習者が自己調整学習能力を身につけることは可能であり、それが効果的な自主学習に結びつくと提案している。たとえ授業課題や単位と直接的な関係のない自主学習においてであっても、学習者の自己調整学習能力の素地を養うことが、効果的に英語学習に取り組める一因と考えて良いだろう。

7

結論

本研究では、情意方略、動機づけ調整方略、学習方略の3つの観点から、英語資格試験のための自主学習における自己調整学習能力について検討した。統計的処理を経てそれぞれの尺度には

一定の信頼性および妥当性があることが認められた。総括すると、自己効力感や目標・学習信念といった動機づけ側面について、さまざまな文脈でその重要性が議論されてきたが、これらは自主学習環境下においても重要な位置づけにあることがわかった。そして、その動機づけの維持や調整のために、考え方や環境に変化をもたらすための方略が用いられるが、同時に、メタ認知的活動を中心とする学習方略が活用されていることがわかった。すべての因子が自主学習頻度および資格試験得点に直接影響を与えているわけではなく、互いに交絡し、最終的に、自主学習の頻度に影響を与える方略と資格試験得点に影響を与える方略に作用することが明らかとなった。

これらの結果は、本研究の参加者の傾向に過ぎず、一般化可能性について議論の余地があることは言うまでもない。特に、参加者の多くは大学1・2年生であり、本格的に留学や就職を意識している学年とは考え方や立場が乖離している可能性もある。同様の尺度を用いた別の参加者集団への実施や、学年別の比較といった追試が必要であろう。また、本研究では、当時全参加者が保持していたTOEICのスコアを資格試験得点としているが、ここに2つの限界点があると考える。1つは、質問項目自体は「4つの英語資格試験」を想定している一方、TOEIC以外の資格試験のスコアは多変量解析に用いることができなかった点である。もう1つは、TOEICのスコアは参加者の自己申告によるもので、参加時期と必ずしも重複しているわけではない点である。受験した試験、受験回数や受験時期は参加者によってさまざまであり、どの時期のスコアを申告し、どの受験の学習方法を想起して回答したかが不明瞭な点で、実施面での限界があったと考えられる。最後に、本研究では学習方略として、英語資格試験の共通科目である語彙・文法・リーディング・リスニングに関する方略を抜粋し、ライティングやスピーキングの学習方略は項目数の増幅と参加者の疲労度を考慮し排除した。そのため、試験別の認知的方略には、本研究とは異なる特徴が見られる可能性がある。今後、自主学習における自己調整学習の一般化とともに、それがどれほど英語力と関わりを持つのか、さらに精緻な尺度を作成する上での課題としたい。

謝 辞

本研究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、助言者である吉田研作先生のご指導により、研究方針が明確に定まり発表に至りました。深く感謝申し上げます。そして、立教大学の高橋里美先生には、親身で建設的なご助言を幾度となく賜りました。心より感謝いたします。また、立教大学の新多了先生の専門的なご示唆およびご助力に、深く感謝いたします。そして、立教大学大学院博士課程在籍の熊谷允岐さん、中野悠稀さんのご協力のおかげで本研究は完遂いたしました。最後に、立教大学大学院生の皆さま、本研究に参加してくださった大学生の皆さま、および、本研究に携わってくださったすべての方々に感謝の意を表します。

注:

- (1) Pintrich et al.(1991) は、あらかじめ「動機づけセクション」と「学習方略セクション」を設けており、分析もそれぞれのセクションに対し別々に実施している。Pintrich et al.(1991) の尺度を参照していることから、本研究もこれに倣って分析を進めた。
- (2) 床(天井)効果とは、ある項目に対する参加者の平均値が極端に低く(高く)、得点分布が偏っていることである(清水・莊島, 2017)。
- (3) ω (オメガ)係数は、因子の項目数に依存せず、各項目の因子に対する負荷量を考慮している点で α 係数よりも正確な信頼性とされている(尾崎・莊島, 2014)。
- (4) GFI (goodness of fit index) および AGFI (adjusted goodness of fit index) は1~0間で推移し、.95以上の値を示すことが望ましい。CFI (comparative fit index) も1.00に近ければ近いほど良いとされ、.95以上が目安とされている。RMSEA (root mean square error of approximation) は .05以下が望ましいとされ、.10以上は不適切と考えられている(田崎, 2015)。

参考文献(*は引用文献)

- * Chang, M. (2005). Applying self-regulated learning strategies in a web-based instruction: An investigation of motivation perception. *Computer Assisted Language Learning*, 18(3), 217-230. <https://doi.org/10.1080/09588220500178939>
- * Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. New York: Routledge.
- * 英語教育2019年6月号(2019). 東京:大修館書店。
- * 英語教育2020年4月号(2020). 東京:大修館書店。
- * Elliot, A. J., & McGregor, H. A. (2001). A 2×2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501-519. <https://psycnet.apa.org/doi/10.1037/0022-3514.80.3.501>
- * Fryer, J. W., & Elliot, A. J. (2008). Self-regulation of achievement goal pursuit. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and application* (pp. 53-75). New York: Routledge.
- * 南風原朝和・市川伸一・下山晴彦(編)(2003).『心理学研究法』東京:放送大学教育振興会
- * Habok, A., & Magyar, A. (2018). Validation of a self-regulated foreign language learning strategy questionnaire through multidimensional modelling, *Frontiers in Psychology*, 9, 1388. <https://doi.org/10.3389/fpsyg.2018.01388>
- * 廣森友人(2015).『英語学習のメカニズム—第二言語習得研究にもとづく効果的な勉強法』東京:大修館書店。
- * 伊藤崇達(2009).『自己調整学習の成立過程—学習方略と動機づけの役割』。京都:北大路書房。
- * Kim, T-Y., & Kim, Y-K. (2014). EFL students' L2 motivational self system and self-regulation: Focusing on elementary and junior high school students in Korea. In K. Csizer & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 87-107). Bristol: Multilingual Matters.
- * Kondo, M., Ishikawa, Y., Smith, C., Sakamoto, K., Shimomura, H., & Wada, N. (2012). Mobile assisted language learning in university EFL courses in Japan: Developing attitudes and skills for self-regulated learning. *ReCALL*, 24(2), 169-187. <http://doi:10.1017/S0958344012000055>
- * Kormos, J., & Csizer, K. (2014). The interaction of motivation, self-regulatory strategies, and autonomous learning behavior in different learner group. *TESOL Quarterly*, 48(2), 275-299. <https://doi.org/10.1002/tesq.129>
- * Lens, W., & Vansteenkiste, M. (2008). Promoting self-regulated learning: A motivational analysis. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-*

参考文献(*は引用文献)

- regulated learning: Theory, research, and application (pp. 141-168). New York: Routledge.
- * Liu, S. H., Lan, Y., & Ho, C. Y. (2014). Exploring the relationship between self-regulated vocabulary learning and web-based collaboration. *Educational Technology & Society*, 17(4), 404-419. <https://www.jstor.org/stable/jeductivechsci.17.4.404>
- * Messick, S. (1995). Validity of psychological assessment. *American Psychologist*, 50, 741-749. <https://doi.org/10.1037/0003-066X.50.9.741>
- * Mills, N. (2014). Self-efficacy in second language acquisition. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA*. (pp. 1-22). Bristol: Multilingual Matters.
- * Mizumoto, A., & Takeuchi, O. (2012). Adaptation and validation of self-regulating capacity in vocabulary learning scale. *Applied Linguistics*, 33(1), 83-91. <https://doi.org/10.1093/aplin/amr044>
- * 新多了(2019).『英語の学び方入門』. 東京:研究社.
- * Oxford, R. L. (2011). *Teaching and researching language learning strategies*. Harlow, UK: Pearson Longman.
- * Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context* (2nd edition). New York: Routledge/Taylor & Francis.
- * 尾崎幸謙・莊島宏二郎(2014).『パーソナリティ心理学のための統計学—構造方程式モデリング』. 東京:誠信書房.
- * Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40. <https://doi.org/10.1037/0022-0663.82.1.33>
- * Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. J. (1991). A manual for the use of the motivated strategies for learning questionnaire (MSLQ). *National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan, Ann Arbor, Michigan, USA*.
- * Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. J. (1993). Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire. *Educational and Psychological Measurement*, 53, 801-813. <https://doi.org/10.1177%2F0013164493053003024>
- * Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475. <https://psycnet.apa.org/doi/10.1006/ceps.1994.1033>
- * Schunk, D. H., & Greene, J. A. (2018). Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed.), (pp. 1-15). New York: Routledge.
- * Seker, M. (2016). The use of self-regulation strategies by foreign language learners and its role in language achievement. *Language Teaching Research*, 20(5), 600-618. <https://doi.org/10.1177%2F1362168815578550>
- * 清水裕士・莊島宏二郎(2017).『社会心理学のための統計学—心理尺度の構成と分析』. 東京:誠信書房.
- * 竹内理(2007).「自ら学ぶ姿勢を身につけるには—自主学習の必要性とその方法を探る」.『Teaching English 8』2-5.
- * 田崎勝也(編) (2015).『コミュニケーション研究のデータ解析』. 京都:ナカニシヤ出版.
- * Tseng, W., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, 27(1), 78-102. <http://doi:10.1093/applin/ami046>
- * 梅本貴豊・稻垣勉(2019).「授業中の動機づけ調整方略と動機づけの変動性の関連」.『心理学研究』90(2), 207-213. <https://doi.org/10.4992/jjpsy.90.18312>
- * 梅本貴豊・田中健史朗(2012).「大学生における動機づけ調整方略」.『パーソナリティ研究』21(2), 138-151. <https://doi.org/10.2132/personality.21.138>
- * Ushioda, E. (2014). Motivational perspectives on the self in SLA: A developmental view. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA*. (pp. 127-141). Bristol: Multilingual Matters.
- * Wang, C., & Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Quarterly*, 51(4), 931-947. <https://doi.org/10.1002/tesq.355>
- * Wang, C., & Schwab, G., Fenn, P., & Chang, M. (2013). Self-Efficacy and self-regulated learning strategies for English language learners: Comparison between Chinese and German college students. *Journal of Educational and Developmental Psychology*, 3(1), 173-191. <http://dx.doi.org/10.5539/jedp.v3n1p173>
- * Wolters, C. A., & Won, S. (2018). Validity and the use of self-report questionnaires to assess self-regulated learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed.), (pp. 307-322). New York: Routledge.
- * Zimmerman, B. J. (1989). Models of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research and practice*. (pp. 1-25). New York: Springer-Verlag.
- * Zimmerman, B. J. (2000). Attaining self-regulation: a social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego: Academic Press.
- * Zimmerman, B. J. (2001). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2
- * Zimmerman, B. J., & Martinez-Pons, M. (1990). Student difference in self-regulated learning: relating Grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51-59. <https://psycnet.apa.org/doi/10.1037/0022-0663.82.1.51>

