EIKEN BULLETIN

vol. 35 2023

第35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35回「英検」研究助成　選考委員

＊は専門選考委員（所属・役職は委嘱当時）　※五十音順

＊和泉　伸一　上智大学 教授

岩崎　紀美子　全日本中学校長会 編集部幹事

＊木村　光男　公益財団法人 日本英語検定協会 専務理事・事務局長

＊小泉　利恵　清泉女子大学 教授

＊斉田　智里　横浜国立大学 教授

志知　芳彦　全国高等学校長協会 英語部会代表・国際教育交流委員会 委員長

＊竹内　理　関西大学 教授

＊寺内　一　高千穂大学 学長

富高　雅代　文部科学省 初等中等教育局 教科調査官

＊西垣　知佳子　千葉大学 教授

博田　英明　全国英語教育研究団体連合会 会長

はじめに

筑波大学　教授
小泉　利恵

日々の英語授業において，指導と評価の一体化が求められている。成績を付けることが主な目的の総括的評価だけでなく，形成的評価として，指導で伸ばそうとした力がどの程度伸びたか，今持つ力の強みと弱みをどのように伸ばしたらよいかの情報を評価から得て，それに基づいて指導内容を修正し，指導後にまた評価を行うというサイクルを継続して行うことが必要とされる。そのような形成的評価の理念は分かるとして，どのように行うべきか，また行っている評価の実践が効果的なのかについて興味や疑問を持つ教員は多いだろう。

しかし，残念ながら教室内で行えるような評価をテーマとする英検研究助成は一部にとどまり，もっと多くの応募があってよいと長年感じている。その背景には，教員養成や研修で評価のトピックが十分に扱われていないこと，そのため評価が研究対象となることや，日本の学校環境での英語評価にも応用できるような先行研究が国内外に多くあることがあまり知られていないこと，さらに日本において，英語の教室内評価を主な専門とする研究者の数が限られ，研究の魅力を十分に発信できていないこと，また数字に強くないと評価の研究はできないなど，研究が難しそうというイメージで躊躇してしまうことなどの理由が考えられる。

しかし，研究がかなり進んでいる分野と比べて，評価の分野には研究して日々の実践に落とし込めるテーマが多くある。例えば，観点別評価に向けて用いるルーブリック（採点基準）はどのようなものがよいか，定期テストや小テストの結果を次の学習につなげるためには，どのような活動や支援を，テストの事前と事後に加えたらよいかなどのテーマは，英語を教える教員ならば考えたことがあるのではないだろうか。また聞くことや小学校での研究など，英検研究助成であまり扱われていない分野と合わせたテーマも考えられる。そのテーマを研究計画にどのように落とし込むかは，申請書のサンプルや，過去の英検研究助成報告書のすべてがオンラインで参照可能なので，参考にしてほしい。

英検研究助成制度は，入賞したら助成金が得られるだけでなく，研究を進め，報告書を作成するにあたって研究助言者から助言を定期的に得ることができる。またその後も研究助言者と交流が続くこともある（私の場合は助言者となった方と今でも連絡を取り合っている）。多くの方からの，特に評価に関する研究の応募を期待している。

Contents

● **はじめに**

筑波大学 教授　小泉　利恵 …… 3

● **報告別講評**　専門選考委員 …… 6

A 研究部門

読解中のノートテイキングが一貫した心的表象に与える影響
― 再話による読解情報の復元を通して ―

研究者：小野 由香子 茨城県／筑波大学大学院　在籍 …… 12

グラフィックオーガナイザーを用いた心的表象構築プロセスの視覚化と評価

研究者：工藤 大奈 東京都／安田学園中学校高等学校　教諭 …… 44
申請時：茨城県／筑波大学大学院　在籍

英字幕付き音声動画の視聴による第二言語の偶発的語彙学習
― 使用する音声動画マテリアルに関わる変数が持つ影響 ―

研究者：黒川 皐月 宮城県／東北大学大学院　学術研究員 …… 56
申請時：東京都／東京外国語大学大学院　在籍

複数テキスト読解でのつながり形成が日本人EFL学習者の統合的理解に与える影響の検証

研究者：三上 洋介 埼玉県／駿河台大学　助教 …… 74
申請時：東京都／東京都立国分寺高等学校　教諭・茨城県／筑波大学大学院　在籍

Contents

B 実践部門

高専生による説明文理解を深める協同的な再話活動
― 状況モデル構築につながる質問項目の比較・検討 ―

研究者：伊東　賢 茨城県／茨城工業高等専門学校　講師・茨城県／筑波大学大学院　在籍 …… 102

日本人高校生に対する，短期間でより効果的なスピーキング活動:
「タスクの繰り返し」の検証

研究者：上野　正和 福岡県／北九州市立高等学校　教諭・茨城県／筑波大学大学院　在籍 …… 131

未知語推測トレーニングの効果:品詞への注目を促す明示的指導を通して

研究代表者：岡田　龍平 群馬県／関東学園大学附属高等学校　教諭 …… 155

中学校の「読むこと」におけるMediation活動を通した思考・判断・表現力の育成

研究者：林　さなえ 兵庫県／兵庫県立芦屋国際中等教育学校　教諭 …… 166
申請時：兵庫県／兵庫県立芦屋国際中等教育学校　教諭・兵庫県／神戸市外国語大学大学院　在籍

AIツールを活用して日本人中高生の学習の主体性を高めることをねらった
ライティング指導

研究者：松井　市子 新潟県／新潟県立津南中等教育学校　教諭 …… 183

C 調査部門

学習者の英語観と情意要因
― グローバル英語に焦点を当てて ―

研究者：池田　幸恵 広島県／広島商船高等専門学校　准教授 …… 206
申請時：広島県／広島商船高等専門学校　講師

● **今回の報告について** 日本英語検定協会 会長・名誉顧問より …… 227

報告別講評

A. 研究部門・報告 I

寺内　一

読解中のノートテイキングが一貫した心的表象に与える影響

― 再話による読解情報の復元を通して ―

【研究者：小野 由香子】

現在の英語教育では，テクストの文法解析ではなく，行間の情報も含めた豊かな読解に基づく思考，判断，表現が求められている。本研究は，再話に役立つノートテイキングの種類を明らかにし，熟達度を考慮した上でそれを効果的に指導するための方法を提案している点で教育的な意義がある。

本研究で取り上げた3種類のノートテイキングは学習者が日常的に行っているものであるため，本研究から得られた示唆は教育現場ですぐに取り入れることができるようになっていることが大きな特長である。統計を活用した量的な分析では，ノートテイキングの種類による再話の差を確認することができなかったが，ノートの質的な分析により，再話の成績が向上するノートテイキングの特徴を明らかにしている。ノートテイキングの種類に差が出なかったことから，特定の方略使用を促すという結論ではなく，学習者が自分やテクストにあった種類を選択し，その上で再話の成績を向上させるためのノートテイキングを指導するという結論につながることも，集団授業による画一的な指導ではなく，個に応じた指導が求められる現在の教育に適合している。

今後，本研究を発展させるためには，実際に実施したノートテイキングの指導が，熟達度を含む個人の特性にどのような影響を受けるのかを明らかにすることが必要であろう。どの(What)種類をどのような(How)方法で，誰に(Who)教えることが効果的なのかがわかれば，教師は目の前にいる学習者に合わせて指導に役立てることができるようになる。

A. 研究部門・報告 II

和泉 伸一

グラフィックオーガナイザーを用いた心的表象構築プロセスの視覚化と評価

【研究者：工藤 大奈】

英文読解をする際に，一つの助けとなるものとして，グラフィックオーガナイザーがある。グラフィックオーガナイザーとは，英語のテキスト情報を図式化して，文章の論理構成や流れを視覚的に提示するものである。工藤氏の研究では，(1)英検で用いられるリーディング問題でグラフィックオーガナイザーがどこまで使用が可能かということと，(2)リーディングをしながらグラフィックオーガナイザーを作成した場合，それがどこまで読み手の読解プロセスと理解を助けることができるかについて調べている。

調査結果は，(1)英検準2級・2級・準1級のとりわけメール文と説明文の内容一致選択問題でグラフィックオーガナイザーが活用可能ではないかと結論づけている。(2)に関しては，設問正答率とグラフィックオーガナイザーの作成との間には，特に強い関係性は見出せなかった。

本研究はそのデザインや分析方法など，今後改善が必要と思われる部分も少なくないが，グラフィックオーガナイザーの役割に注目し，それがどこまで使用可能で，どれほど読解過程で役立つものかを調べた視点は重要である。英語授業での読解の支援策として，語彙指導や構文解説，あるいは日本語訳を与えたりすることばかりに終始してしまうことがある。しかし，グラフィックオーガナイザーのような視覚情報を与えたり，作成させたりすることも，学習者主体の一つの有益な支援方法となり得るだろう。それがどういったテキストに対して有効で，どのような効果があるかについては，今後も研究を続けていくことが重要だろう。グラフィックオーガナイザーの示し方や作成手順も含めて，更なる研究が進むことを期待する。

A. 研究部門・報告Ⅲ

西垣 知佳子

英字幕付き音声動画の視聴による第二言語の偶発的語彙学習

― 使用する音声動画マテリアルに関わる変数が持つ影響 ―

【研究者：黒川 皐月】

本研究の目的は，英語あるいは日本語の字幕付きの英語の動画の視聴が，偶発的語彙習得に与える影響を調査することであった。参加者は，CEFRのB1からB2の，2,000語から3,000語レベルの語彙知識を持つ大学生と大学院生148名であった。動画はTED-Edから選定され，動画の語彙レベルと参加者の語彙レベルが綿密に統制された。参加者は動画を2回視聴し，日本語字幕と英語字幕の有無の組み合わせにより4群が設定された。さらに2群が追加され，画像のみから得られる内容理解，意図的な語彙学習の効果が検証された。効果検証には，意味理解の深さが異なる3種類のターゲット語の語彙テストが，指導前と指導終了の約1週間後に実施された。

収集されたデータに対して，一般化線形混合モデルのロジスティック回帰分析を行った結果，英語の動画を学習教材にする際は，動画に対する学習者の語彙カバー率に注意する必要があることが示唆された。具体的には，動画は映像・音声・音響が伴うため，カバー率が90%で充分な内容理解がなされ（読解では95%），95%になると未知語に注意を向ける余裕が生まれ，日本語，英語の字幕が無くても，偶発的な語彙習得が起こった。本研究の結果を踏まえると，動画によって発音，場面・状況を伴った語彙習得が引き出されることから，カバー率95%以上の動画を使った偶発的な語彙習得を加えることで，教室での教科書による意図的な語彙学習を補完できると言う貴重な教育的示唆が得られた。

以上から本研究は，動画を語彙学習教材として利用する際の教材選定に有効な指標を示した点で，動画の活用，また，偶発的な語彙習得に関する貴重な基礎的研究と言える。この領域の研究は日本では多くない。今後も発展的に研究を継続していただくことを願っている。

A. 研究部門・報告Ⅳ

小泉 利恵

複数テキスト読解でのつながり形成が日本人EFL学習者の統合的理解に与える影響の検証

【研究者：三上 洋介】

複数のテキストを読み進めて情報を得ることは，日本語ではよくあり，英語でも増えつつある。またテストにも，複数テキスト読解が取り入れられつつあり，例えば，令和7年度大学入学共通テストの英語（リーディング）試作問題や，令和5年度全国学力・学習状況調査の調査問題の小学校・中学校（国語）で確認できる。三上氏は，そのように現実先行的に広がっている複数テキスト読解の中で，テキスト間のつながりを，大学生がどのように形成し，統合的に理解しているかを検証した。テキスト間の関係も，内容の方向性が同じ「相補的テキスト」と異なる「対立的テキスト」を読んだ場合の相違など，母語で行われた先行研究を日本の英語教育の場面で応用するための基礎研究として実りある報告を行っている。

また論文では，読者にとって手順が分かり，再現性研究ができるように丁寧な記述がされている。例えば資料1や2を見ることで，採点基準やプロセスの理解が深まるだろう。透明性の高い記述は，研究者としての誠実さを示すものであり，また該当分野への興味と，一般化と文脈化に向けた研究と実践への関心を高める重要なステップである。

本テーマは英語教員だけでなく，テストや教材作成の団体も関心があるテーマであろう。関係者の関心を拾い上げ，相補的・対立的テキストにとどまらない複雑なテキストの読解や，その過程で鍛えられる批判的思考力の特徴など、多方面への研究・実践として発展していくことを期待する。

B. 実践部門・報告Ⅰ

西垣 知佳子

高専生による説明文理解を深める協同的な再話活動

— 状況モデル構築につながる質問項目の比較・検討 —

【研究者：伊東 賢】

本研究は，昨今，英語の授業改善の方策の1つとして注目される「再話」に関する研究である。再話には，学習者の英語レベル，教材の種類と難易度，支援の有無とその種類，再話の際の使用言語（英語・日本語）等，複雑な要因が関与する。本研究では，CEFRのA1～A2レベルの国立高等専門学校2年生が，英語の説明文を学び，再話に5回取り組んだ。参加者140名を，英語による再話の有無，再話が止まった時の再話再開の足場がけ質問の有無という条件の違いで3群に分け，発話の流暢性の向上，また，文面どおりの内容理解と推論を伴う内容理解（状況モデルレベルの理解）の効果を検証した。

その結果，英語を使った再話では，発話の流暢性は，発話スピードよりも発話量の増加，ポーズの長さの短縮に効果があった。また，推論を伴う内容理解よりも，文面どおりの内容理解のほうに効果があり，学習者はテキストの英語表現に注意を向けていたことが推測された。なお，今回は，聞き手から再話者への足場がけ質問が単語レベルであったため，2つのレベルの内容理解に効果が見られなかった。初級学習者が，英語の再話で内容理解を深めるには，足場がけ質問の出し方を練習する必要があることも推測された。

本研究を通して，英語による再話が引き出す効果が示された。また，初級学習者の英語の再話では，どのような認知処理を経て内容理解が進むのか，あるいは停滞するのか，その学習過程の可能性が示され，どのような活動を行うと，どのような効果を期待できるのか，また，どのような練習が効果をあげるのかに関する基礎データを得られたことは意義深い。伊東氏は継続して再話研究に取り組んでおられることから，今後の研究の進展に多いに期待する。

B. 実践部門・報告Ⅱ

小泉 利恵

日本人高校生に対する，短期間でより効果的なスピーキング活動：「タスクの繰り返し」の検証

【研究者：上野 正和】

現在の学習指導要領に基づき，高校では様々なスピーキング活動が行われている。その中で上野氏は，週3回2週間，集中的に絵描写タスクの繰り返しを行い，その流暢さへの効果を検証した。その結果，流暢さの中で伸びる要素を示し，また方法ではインターリーブ型を継続する形がより効果的であることを導いている。また絵描写の繰り返しの間には，模範解答の提示や振り返り活動を行い，毎回のペアを変えるなどの教育的支援も取り入れ，生徒が話すことへの自信を高めたことが報告されている。

研究として要因の特定のためには，要因を絞り，それ以外の要因は入らないような工夫が通常必要とされる。しかし，通常の英語授業の中で行う研究では，それが厳密に行えなかったり，また行うべきではなかったりする。今回は教育的支援を重視した形で行い，意義ある研究になった。しかし研究結果の解釈としては，タスクの繰り返しとその後の支援を含めた効果と理解するべきだろう。また上野氏は，本研究の手順や資料を詳細にまとめているため，非常に参考になる。さらに言えば，著作権を踏まえつつ，使用した絵や資料のファイルなどを，できるだけ多く補助資料として公開することを望む（公開例：https://osf.io/bjtg6/）。

今後はこの研究を土台として，5～10分程度の帯活動でタスクの繰り返しを取り入れた場合の効果や，絵描写後に生徒同士のやり取りを入れた場合の効果などの実践や研究が広がっていくことを祈る。

B. 実践部門・報告Ⅲ

寺内　一

未知語推測トレーニングの効果：品詞への注目を促す明示的指導を通して

【研究代表者：岡田 龍平】

現実の言語使用場面において，未知語の推測は頻繁に行われる処理であるため，推測能力を育成する指導の効果を明らかにすることは，実践的な言語使用につながる点で大いに価値がある。

特に，未知語の推測能力の指導方法はいまだ確立されていない中で，著者は形式重視のトレーニングと意味重視のトレーニングを実施している。これらは教師が日々の指導に組み込むことが容易であるため，どちらがより効果的かを明らかにする本研究は教育的な意義が大きい。統計的な分析では，トレーニング方法の違いによる指導効果の差は明らかにならなかったが，トレーニングの前後で品詞を同定する能力の向上が見られている。このことから，筆者が取り上げた2つのトレーニングは一定の教育効果があることが証明された。

なお，推測の手がかりを同定する能力と未知語の意味を推測する能力について，本研究ではトレーニングの効果が確認できなかった。著者はこの理由として，指導時間や練習回数を挙げている。もちろん，トレーニングの量を増やせば効果が上がることは期待できる。ただし，未知語の推測プロセスは非常に複雑であり，未知語の種類や推測に生かせる背景知識や文脈も様々に異なるため，実際にはどれか1つの指導が効果的というよりは，複数の指導法を組み合わせることが最も効果的であることが推察される。したがって，今後の研究として，学習者のレベルや英語学習の目的等に合わせて，どのような指導法の組み合わせや指導順序が効果的なのかを明らかにする研究も必要となってくるであろう。

B. 実践部門・報告Ⅳ

竹内　理

中学校の「読むこと」におけるMediation活動を通した思考・判断・表現力の育成

【研究者：林 さなえ】

本実践は，英語の「読むこと」において，Mediation活動を通して，中学生の思考力・判断力・表現力を高めることを意図している。具体的には，中学3年生20名が，ペアワークの形態で，別のペアとは違う内容の物語文・説明文を読み，その内容を相手に伝える活動を行った。この中で，聞き手には話を聞きながら内容に関する質問を行うことを，また話し手には読んだ内容をもとにそれらの質問に答えることが求められた。この実践の結果，生徒たちのリーディング教材の理解度は，テキストのジャンルに関わらず上昇した。また，内容に関する質問や，推論を活用した質問など，生徒が内容理解を深めるために行ったやり取りの過程から，思考力・判断力・表現力を高めるという目的に関しても，一定の成果が得られたことが確認された。

MediationというCEFR（ヨーロッパ言語共通参照枠）にも位置づけられている概念を具体的な活動として取り扱い，思考力・判断力・表現力という測定の難しい課題に，質と量の両面から果敢に挑んだことは高い評価に値する。今後この実践が，単発ではなく長期にわたり行われ，その成果がどのように言語能力の向上にも役立つかという問題まで踏み込むことが出来れば，中等教育における英語教育に大いに貢献するものと期待できる。

B. 実践部門・報告Ⅴ

斉田 智里

AIツールを活用して日本人中高生の学習の主体性を高めることをねらったライティング指導

【研究者：松井 市子】

AI技術の進化をどのように英語教育に活用していくかは，教員にとって大きな関心事であり，実践研究の発展が期待される分野でもある。松井氏の問題意識は，ライティングの指導にAI技術を取り入れることで，生徒はより多くのアウトプットの機会を得ることができ，教師はフィードバックの負担を減らすことができ，結果として学習の主体性を高めることができるのか，という点にある。

本稿は，公立の中等教育学校全6学年の生徒を対象に実践研究を行った報告である。教師の指導は介さず，AIツールによるフィードバックのみを頼りに，中高生は自力でライティング力を向上させることができるのか，またこの実践を通して主体的に学習に取り組む態度が育成されるのか，という研究課題の下，ライティングの支援ツールとして3種類のソフトウェアを生徒に紹介し，生徒自身に最適なものを選ばせ，継続的にタスクに取り組ませた。生徒のライティングにおけるCEFRレベルや自律学習能力スコアなどに，事前と事後・遅延で違いがみられるかを生徒からのアンケート回答をもとに分析し，教員からも意見を聴取した。GIGAスクール構想，AIツール，個別最適な学び，主体性といった現代的な教育テーマを英語ライティングの指導に適用した意欲的な実践研究である。生成AIの登場により，教育分野におけるAI活用に関する議論が一層活発に行われている。今後のこの分野の研究の発展に注目したい。

C. 調査部門・報告Ⅰ

斉田 智里

学習者の英語観と情意要員

― グローバル英語に焦点を当てて ―

【研究者：池田 幸恵】

英語使用が拡大する中，英語の概念が母語話者英語から非母語話者の多様な英語変種へと変化してきている。英語の多様性を重視する「グローバル英語」（Global Englishes）という新たな英語概念が浸透する中で，英語学習者の英語観や，動機付け，英語完全主義，スピーキング不安といった情意要因も変化していると考えられる。池田氏は，この点を明らかにするために，高専生を対象に，英語学習観と情意要因との関係を調査した。高専生は，実践的技術者を目指し，アジア圏を中心とする非英語圏の人々との国際交流の機会が多く，英語概念の変化が彼らの英語学習観や情意面にどのような影響を与えているかを調査することは，この研究分野に新たな知見をもたらし，教育的示唆が期待される。

池田氏は，先行研究で使用された英語学習観と情意要因に関する質問紙を改良してデータを収集し，新たに因子分析を行い，相関分析と重回帰分析の結果をもとにパス解析を行い，丹念に変数間の関係解明を試みている。複雑なパス解析による結果の解釈は難しい面もあるが，高専生の英語教育においては，グローバル英語志向と日本人英語話者としてのアイデンティティを育むことが有用であるとの結論に至った点は，グローバル英語の時代の英語教育の方向性を示したものとして意義が大きい。池田氏も指摘するように，高専生の英語観が5年間の教育課程でどのように変化していくのか，英語観の変容に最も影響を与える要因は何か，といった長期的かつ縦断的な研究が今後期待される。

● **第35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	2022年2月1日～4月30日
選　考	2022年5月26日～7月12日
助成金贈呈式	2022年8月18日
研究期間	2022年7月～2023年8月
報告書提出	2023年8月

A 研究部門 ● 報告Ⅰ ● 英語能力テストに関する研究

読解中のノートテイキングが一貫した心的表象に与える影響

― 再話による読解情報の復元を通して ―

研究者:小野 由香子　茨城県／筑波大学大学院　在籍

《研究助言者:寺内　一》

概要　本研究では，読解中の活動であるノートテイキングがテキスト内容の記憶にどのような影響を与えるのかを検証することを目的に，アンダーライン/ キーワード抽出/ 図式化のいずれかの読解タスクを与え，読解後の再話プロトコルを比較した。二元配置分散分析の結果，各タスク条件で再話の成績に有意性は見られなかったが，熟達度の主効果は有意であることが明らかになった。また，非テキスト情報について，熟達度上位群の方が有意にSUBの出現率が高かったが，ADDやSUMについては熟達度間で差はなく，各タスク条件間でも有意な差は見られなかった。さらに，協力者のノートと再話プロトコルの質的な分析から，ノートの取り方と再話成績には関係がある可能性が示唆された。本研究から，生徒にノートテイキングを指導する際には，生徒の熟達度を考慮し，英文の構造がわかるようなノートをとることを伝えることで，より強固な心的表象を構築できることが示唆された。

1 はじめに

1.1 研究背景

日本の教育現場では，グローバル化した世界における英語力の重要性がますます認識されるようになっている。英語を国際的なコミュニケーションの重要なツールとして捉え，日本の教育システムは英語教育へのアプローチを積極的に改良してきた。従来の方法では文法や語彙の習得に重点が置かれることが多かったが，現代の教育学の枠組みでは，読解力，批判的思考，効果的なコミュニケーションなど，実践的な言語能力の育成が優先されている。このような進化する教育状況の中で，ノートテイキングは読解力や記憶保持と交差する重要な認知スキルとして浮上してきた。ノートテイキングは単なる機械的な活動ではなく，テキスト情報に積極的に関与し，重要なポイントを選択し，後で見直すために整理する方略的なアプローチである。しかし，ノートを取る方法の有効性や，それらが記憶の想起に与える影響は，依然として複雑で多面的な研究分野である。先行研究では，認知方略としてのノートテイキングの重要性は認められているものの，第二言語学習者を対象に，様々なノートテイキング方略が読解熟達度やその後の理解力とどのように相互作用するかを検証した包括的な研究は乏しい。これらの要因の相互作用を理解することは，教育者が多様な学習者のニーズに応え，英語教室における読解力を高める教育的介入をデザインする上で不可欠である。よって，本研究では，ノートの取り方，熟達度，再話成績の相互関係を調査することで，英語学習者の読解力と記憶保持力を高める教育実践に貢献することを目的とする。

1.2 EFL学習者の英文読解

読み手があるテキストを読んで記憶に残ったものを心的表象(mental representations)といい，テキストの理解とは，読み手がそのテキストに即して首尾一貫性のある心的表象を構築した状態を示す(門田・野呂，2001)。心的表象には3つのレベルがあり，それぞれ表層的記憶(surface memory)・命題的テキストベース(propositional textbase)・状況モデル(situation model)と呼ばれている。第二言語読解研究では，テキストのより深い理解のためには学習者が適切な状況モデルを構築することが大切であると考えられており，状況モデルが強いほど，テキスト記憶を使用する必要がある後のタスク(例:想起)において，より良いパフォーマンスが可能になる(Horiba, 2013)。状況モデルを構築するためには，読み手が背景知識を活性化させ，推論を生成する必要があるが(卯城，2009)，読み手の熟達度が低い場合やテキストが難しい場合，語の意味や文法解析などの下位レベルの処理に認知資源を費やしてしまい，意味解釈や推論などの高次の処理を行う余裕がないと考えられている(門田・野呂，2001)。よって，教育現場では，そのような学習者が文間・段落間の結び付きや関係性に着目し，文章全体の意味を捉えることができるような手立てを講じる必要がある(市澤・吉野，2018)。意味のある学習を達成するためには，学習者が積極的に情報処理に関与する必要があり，この情報処理への能動的な関与を促進する1つの方法として，学習者の学習方略を促進することが挙げられる。

1.3 読解方略

学習の効率を高めるため，学習者は様々な学習方略を活用している。読解中に使用される方略として，(1)暗記，(2)構成，(3)精緻化，(4)モニターなどが挙げられ，それぞれ (1)文中に線を引いたり情報を書かれているとおりに繰り返したりする方略，(2)テキスト内の情報を統合したり，情報どうしを関連付けたり，並べ替えたりする方略，(3)文章内の情報と文章外の情報を結び付ける方略，(4)自己の学習状況を評価する方略を指す(卯城，2009)。

読解方略の1つであるノートテイキングは学習者がノート・配布資料・テキストにメモしたり下線を引いたりする活動を示し(小林，2000)，学生の間で非常に一般的な習慣である(Hedgcock & Ferris 2009)。読解中のノートテイキングはテキストに対する注意力の向上や，記憶の強化に役立つとされ(Faculty, 2006)，理解，選択，作成といういくつかのプロセスを伴うため，深い理解を促す(Williams & Eggert, 2002)と考えられている。

1.4 読解におけるノートテイキング

ノートテイキングが学習に与える効果については，大きく2つの仮説に分類される(Di Vesta & Gray, 1972)。1つはencoding hypothesis(e.g., Kiewra et al., 1987)というもので，この仮説はノートを記録するプロセスが学習を促進するという考えに基づいており，生徒の注意力を高めること，文章の構成に対する意識を高めること，情報を記憶に保存すること，学習者が以前に学習した情報と教材を比較することを促すことなどが含まれる(Rahmani & Sadeghi, 2011)。もう1つはexternal storage hypothesisとよばれるもので，復習やリハーサルのためにノートを利用できるようにすることで，長期的な利益をもたらすと考えられている。本研究では，読解中のノートテイキングとして，encoding hypothesisに焦点を当てる。

ノートテイキングがテキスト読解に与える影響については，過去の研究によって様々な観点から探究されてきた。例えば，Puji(2019)は，インドネシアの高校生40名を対象とし，ノートテイキングを行ったグループと行わなかったグループで事前テストと事後テストの結果を比較した。テストはTOEICのリーディング問題が使用された。その結果，ノートテイキングを行ったグループは，事後テストで有意に得点が向

上していることが明らかになった。また，Bahrami & Nosratzadeh（2017）は，イラン人EFL学習者を対象に，ノートの取り方の指導を与えたグループとそうではないグループで，テストの結果に違いがあるか調査した。テストは多肢選択式課題で構成され，協力者全員が事前テスト，事後テストに回答した。その結果，ノートの取り方を指導したグループでは，事前テストと事後テストの平均点に有意に差が見られ，ノートテイキングは読解力向上に有効であることが明らかになった。

ノートテイキングが読解を促進することを主張する先行研究によれば，ノートテイキングの利点は，ノートテイキングが学習者の記憶保持を助け，文章中の情報のつながりを作るのに役立っているためと考えられている（Chang & Ku, 2015）。Cook & Mayer（1983）によれば，学習者はノートを取る過程でテキスト情報を深く分析し，主要な概念と二次的なアイデアを区別し，読解内容と関連する予備知識を有意義に統合することができるようになるため，ノートを取ることは，学習者がテキストの首尾一貫したアウトラインを作成し，情報をよりよく整理するのに役立つとされている。

しかし，ノートを取るグループと取らないグループの学習成果を比較したメタ分析（e.g., Kobayashi, 2005）では，ノートテイキングの平均効果量は0.22であり，エンコーディング効果が期待されるほど高くないことが示されている。これは，ノートの取り方や学習成果測定法の種類によって，どの程度生成的処理が活性化するかが異なるためであると考えられている。すべての学習者が効果的にノートをとれるわけではないため，講義やテキストに含まれる重要なアイデアのごく一部しかメモできず，ノートテイキングの利点を享受できない可能性がある。先行研究では，読み手はある目的をもって読解を進めており，熟達度の高い読み手は目的をきちんと把握し，目的に応じて文章処理の方略を柔軟に変化させるが（Pressley & Afflerbach, 1995），熟達度の低い読み手は目的をうまく把握できず，重要な情報とそうでない情報の取捨選択に困難を感じる可能性があると考えられている（卯城, 2009）。よって，読解前に方略の指示を与えることで，学習者が一貫した心的表象を構築できるようになる可能性がある。

ノートテイキングの指導について，Rahmani & Sadeghi（2011）は，イランのEFL大学生108人を対象に，ノートテイキングの指導と読解への影響を調査した。実験群はグラフィックオーガナイザーをガイドとしてノートの取り方に関するトレーニングを受け，統制群は何も指導を受けなかった。二元配置分散分析の結果，理解度テストと想起テストの両方で，実験群の成績が有意に良いことが示唆された。また，筆記による想起の分析でも，実験群がより重要なアイデアを記憶しており，アイデア間の関係をよりよく識別していることが示された。さらに，ノートテイキングでは，実験群の方が対照群よりも多くの単語や重要なアイデアをメモしていることが明らかになった。Rahmani & Sadeghi（2011）の研究結果と同様に，ノートに含まれる単語や概念の数は，テストの成績と正の相関があることが他の先行研究でも明らかになっている（e.g., Boyle, 2010）。

1.5 ノートテイキングの種類

以上のように，先行研究ではノートテイキングの有用性や指導の効果が検討されてきたが，第二言語学習者を対象に，ノートテイキングの種類が読解成果に及ぼす影響について検討した研究はあまり見られない。ノートテイキングにはアンダーライン，重要な情報の抽出やパラフレーズ，図表の作成など様々な種類があり（齋藤・源田, 2008），先行研究では，テキストが難解かつ読解時間が十分に与えられている場合，読解時に下線を引く行為が有効であることや（魚崎 ・伊藤 ・野嶋, 2003），キーワードの強調が，強調した部分の再生成績を高めること（関, 1997）などが明らかになっている。また，テキスト内容の図式化は，二重符号化理論（e.g., Paivio, 1986）の観点から有用であると考えられており，同じ情報が言語と視覚の2つのフォーマットで符号化されることで，2つのモードが相対的に関与し，長期記憶から概念を容易に取り出すことができる（Robinson & Molina, 2002）。テキスト理解を助ける図のうち，英文内の概念の関係や構造を図示するものはグラフィックオーガナイザー（GO）と呼ばれており，第二言語研究において，GOは文章理解に役立つことが明らかになっている（e.g., Suzuki, 2007）。

複数のノートテイキングを比較した数少ない研究のうち，Lonka, et al.(1994)は，200人の大学入学志願者のノートテイキング方略(アンダーラインとコンセプトマップ)が，テキストの詳細な学習，総合的課題，批判的分析にどのように影響するかを評価した。協力者は，本文中または別の紙にノートをとることができた。その結果，アンダーラインは詳細な学習と総合的な課題を改善し，コンセプト・マッピングは内容の批判的分析に影響を与えることが明らかになった。また，Chang & Ku(2015)によれば，ノートを取る際，要約・概念マッピングなど，情報を解釈し，予備知識を取り入れながら自分の言葉で書き直す際にはテキストの理解力を高める(Slotte & Lonka, 1999)が，テキストを逐語的にコピーしてノートを取ることは，情報のより深い処理を妨げる可能性がある。これらの先行研究を踏まえると，効果的なノートテイキングの指導を行うためには，どのようなノートを取ることが読解を促進させるのか，ノートテイキングの種類やノートの取り方について調査する必要がある。

1.6 先行研究の限界点

先行研究では，読解中のノートテイキングが内容理解に有用であることや(e.g., Puji, 2019)，読解力の向上に役立つことが示されているが(e.g., Bahrami & Nosratzadeh, 2017)，多肢選択式課題などの内容理解テストの結果が着目され，読解中に構築した心的表象を読解後にどの程度思い出せるか検証した実験はほとんど見られない。また，効果的なノートテイキングを行うためには指導が重要であるが，読解前に指示するノートテイキングの種類によって読解後の心的表象が異なるのかを検証した先行研究は限られており，読解中のタスクとしてのノートテイキングが，内容理解・記憶にどのような影響を与えるのか調査する必要がある。本研究ではこの限界点を解決するために再話に着目し，読解後の情報復元に焦点を当てる。また，読み手の読解熟達度によって，ノートテイキングの質や量に違いが見られる可能性があることから，熟達度を考慮し，読解タスクとしてのノートテイキングが読解後の心的表象に与える影響を明らかにする。

2 方法

2.1 目的

本研究では，読解中の活動であるノートテイキングがテキスト内容の記憶にどのような影響を与えるのかを検証することを目的に，3つの研究課題(RQ)を立てた。

RQ1 ノートテイキングの種類は再話成績に影響を与えるか

RQ2 熟達度はRQ1の結果に影響を与えるか

RQ3 どのようなノートの取り方が再話を促進させるか

2.2 方法

2.2.1. 協力者

筑波大学に所属する大学生・大学院生36名(男性19名，女性17名)。年齢幅は18歳～32歳(M = 21.9, SD = 2.48)で，専攻は人文，教育，心理，物理，看護など様々であった。日本語を母語とするEFL学習者で，海外経験が1年未満である協力者を対象とした。協力者は，アンダーライン/ キーワード抽出/ 図式化のいずれかの条件にランダムに振り分けられた。

2.2.2. マテリアル

(1)熟達度テスト:協力者のL2読解能力を測るためのテスト。英検の過去問(2級, 準1級)から20問の長文読解問題を出題した(表1)。信頼性係数を示すCronbach α は α = 0.68で, 中程度の信頼性が得られた。

■表1: 熟達度テストの出題級とトピック

大問	出題級	トピック
1	2級	An Extraordinary Machine
2	2級	The Evolution of Laughter
3	2級	Desert Delight
4	準1級	Competing against Braille
5	準1級	The British in India

Note. 例文は巻末資料参照

(2)読解テキスト:Tilstra & McMaster(2013)で使用されていた, 難易度が同程度の2つの説明文テキスト(Animal Talk, The Rodeo)を採用した(表2)。これらのテキストはEFLを対象とした研究(e.g., Kimura, 2022)でも使用されており, テキストが長すぎず, なじみのないトピックであることなどの要件を満たしていた。Word Level Checkerを基に難しい単語には注釈をつけた。

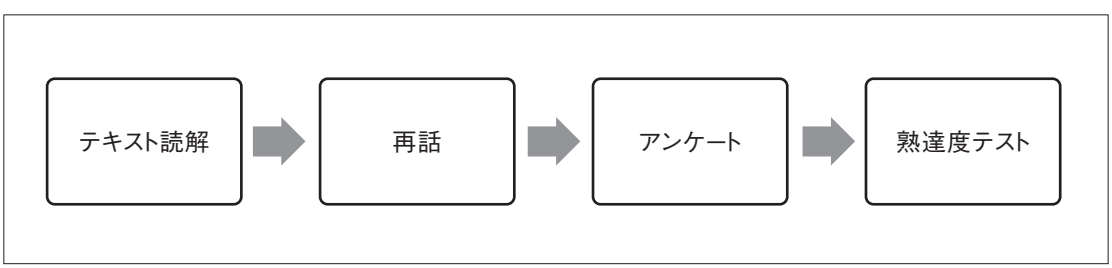
■表2: 読解テキスト

Text	Title	Words	Sentences	FRE	FKGL
A	Animal Talk	269	19	74.3	6.3
B	The Rodeo	305	20	74.7	6.5

Note. テキスト内容は巻末資料参照

2.2.3. 手順

調査ではZoomを使用し, オンラインを通して個別で実施した。協力者はまず実験用のGoogle Formを開き, 実験内容に関する説明に目を通したあと, 基本情報を入力した。入力後, 口頭で指示が与えられ, 読解中に行うタスク(ノートテイキング)の説明, 読解時間に制限がないこと, 読解後, テキスト内容に関する内容理解タスクを行うことなどが伝えられた。協力者はテキストを読みながら各条件でノートテイキングを行い, 読解終了後, Google Form上で内容理解タスクに回答した。内容理解タスク完了後, テキストの難易度や指示されたノートテイキングが役に立ったかどうかなどについて問うアンケートを実施した。実験の後半では, 協力者のL2読解能力を測るための熟達度テストを実施した。問題はPDFで配布し, 回答はGoogle Formで収集された。全体の実施時間は60分〜75分程度であった。



■図1: 手順

2.2.4. 採点

(1)熟達度テスト

1問1点の20点満点で採点し，テストの信頼性係数を算出した。

(2)再話

Ikeno(1996)に基づくIdea Unit(以下IU)による採点と，再話分類法(Kucer, 2009; Kucer & Silva, 1999)による採点の2つを使用し，採点を行った。2名の評価者でテキストをIUに分割し，不一致箇所については協議を行った。評価者間の一致率を検討するため，カッパ係数を求めた結果，テキストA(Cohen's k = .836, p < .001)テキストB(Cohen's k = .88, p < .001)ともに高い水準で一致していることが明らかになった。

IUによる採点では，再話プロトコルの中で，1つのIUの情報の3分の2が再現されていた場合，1点を与えた。再話プロトコルのうち，30%を独立して採点し，不一致箇所については協議を行った。IUによる採点で1点を与えなかった，文字通りのIUの再生ではないプロトコルは，再話分類法を基に非テキスト情報として分類した。再話分類法による評価も同様に，再話プロトコルの30%を2名の評価者で独立して評価し，再話のプロトコルは可能な限りIUと照合した。各再話プロトコルを後述のカテゴリーに分類した後，不一致箇所について協議を行った。

■表3: 再話分類法(Kucer, 2009; Kucer and Silva, 1999)に基づくカテゴリー

Category		Category
1	Substitution	再話で表現されたアイデアが本文中のアイデアの置換である。意味的に受け入れられ，本文で表現されたアイデアの修正を表すもの。
2	Addition	再話で表現されたアイデアはテキストにはないが，意味的に許容できる。暗黙のテキストの意味や，実行可能な推論を表すことがある。
3	Summary	本文中の少なくとも2つの別々のアイデアが，再話では1つの一般的アイデアに凝縮されている。

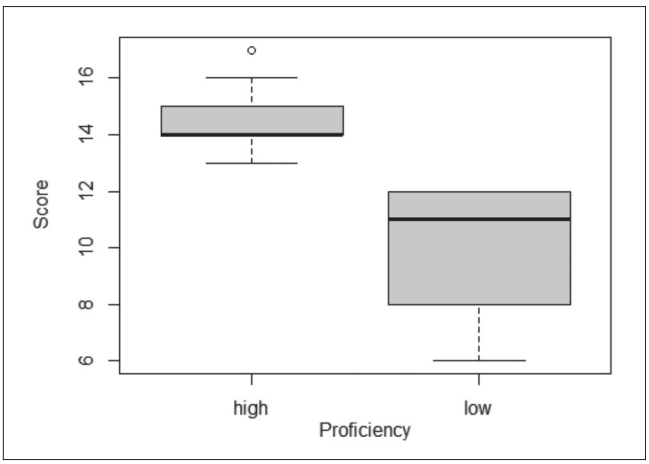
3 結果

3.1 熟達度テスト

平均値により，協力者を上位群/下位群に分類した。それぞれの群の得点についてt検定を行った結果，$t(29) = 8.26$, $p < .001$, $d = 2.68$ [95%CI = 1.75, 3.61]で，上位群の方が有意に得点が高くなっていることが明らかになった。

■表4: 全体のスコア

Proficiency	*n*	*M*	*SD*	*Min*	*Max*
High	17	14.6	1.19	13	17
Low	19	9.8	2.08	6	12



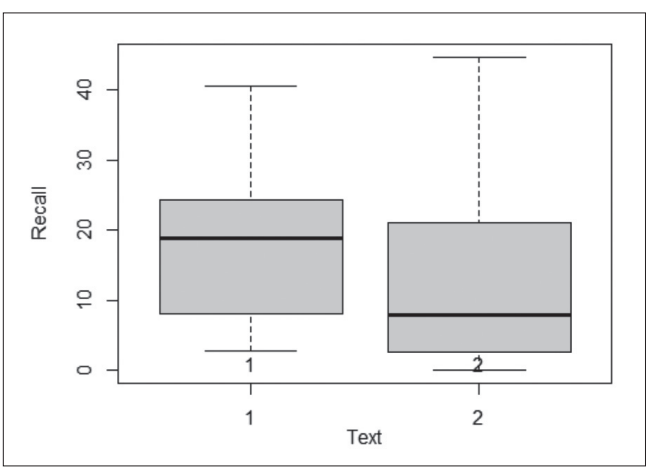
■図2: 熟達度別の再話得点(箱ひげ図)

3.2 再話プロトコルの再生率(IU)

実験では，テキストの内容が再話の結果に影響を与えることを防ぐため，同レベルの2種類のテキストを使用した。2種類のテキストにおける再話の結果について*t*検定を行った結果，$t(31) = .88$, $p = .385$, $d = 0.3$ [95%CI = -0.38, 0.98] で，テキスト間で差はないことが明らかになった。

■表5: 再話の再生率(IU)

Text	*n*	*M*	*SD*	*Min*	*Max*
1	19	17.05	10.51	2.7	40.5
2	17	13.46	12.94	0	44.7



■図3: テキスト別の再話得点(箱ひげ図)

協力者の熟達度と再話の成績(%)との関連について調査するため，2-way ANOVA [2(熟達度:高vs低)×3[タスク条件:アンダーライン，キーワード抽出，図式化]を実施した。その結果，熟達度の主効果は$F(1, 30) = 65.46$, $p < .001$, $\eta_G^2 = .651$で有意であった。タスクの主効果は$F(2, 30) = 1.33$, $p = .281$, $\eta_G^2 = .026$で有意ではなかった。また，熟達度とタスクの交互作用は$F(2, 30) = 1.24$, $p = .305$, $\eta_G^2 = .025$であり，有意ではなかった。

■表6: 熟達度別の再話得点(%)

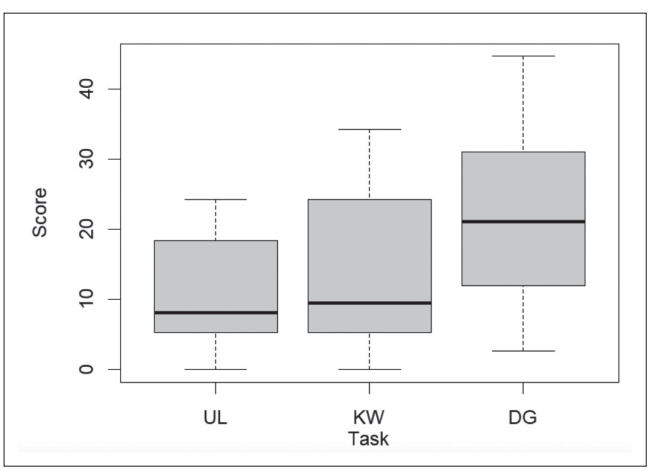
	Proficiency	Task	*n*	*M*	*SD*
Retelling Score(IU)	High	UL	5	14.92	9.00
		KW	7	13.33	10.97
		DG	5	20.64	14.88
	Low	UL	8	8.03	6.52
		KW	5	14.96	11.2
		DG	6	23.8	10.89

Note. UL: アンダーライン条件, KW: キーワード条件, DG: 図式化条件

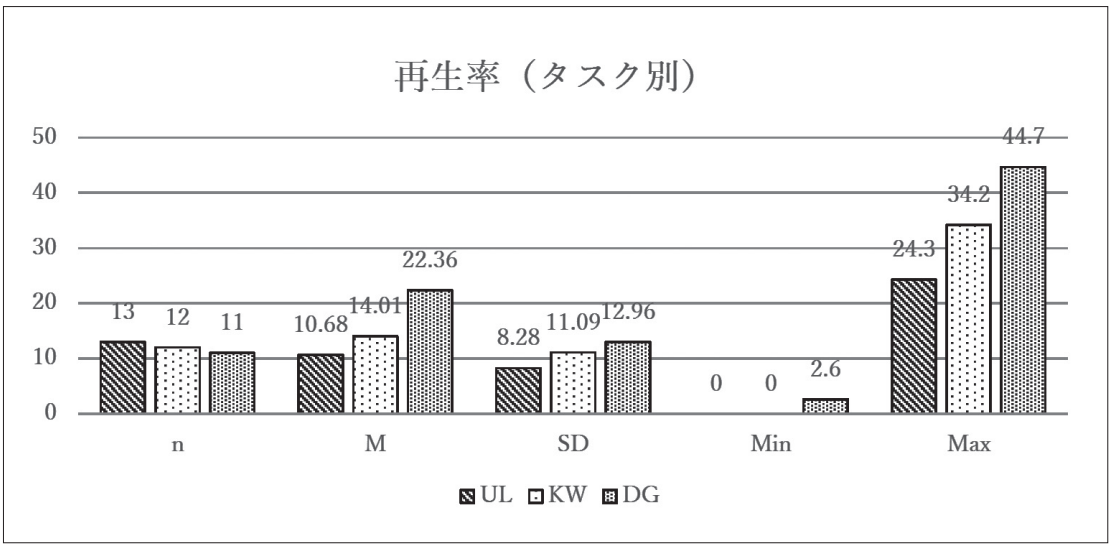
次に, タスク別の再話得点について, 対応なしの一元配置分散分析を行った。その結果, $F(2, 33) = 3.31$, $p < .05$, $\eta_G^2 = .167$となり, 5%水準で有意であった。そこで, Shafferを用いて多重比較を行ったところ, キーワード条件と図式化条件, アンダーライン条件とキーワード条件の間には有意差は認められなかったが, アンダーライン条件と図式化条件の間では, 図式化条件の方が有意に得点が高かった($p = .017$)。

■表7: タスク別の再話得点(%)

Task	*n*	*M*	*SD*	*Min*	*Max*
UL	13	10.68	8.28	0	24.3
KW	12	14.01	11.09	0	34.2
DG	11	22.36	12.96	2.6	44.7



■図4: タスク別の再話得点(箱ひげ図)



■図5：タスク別の再話得点（%）

3.3 再話分類

熟達度別に，非テキスト情報の出現率を計算した。結果を表8に示す。

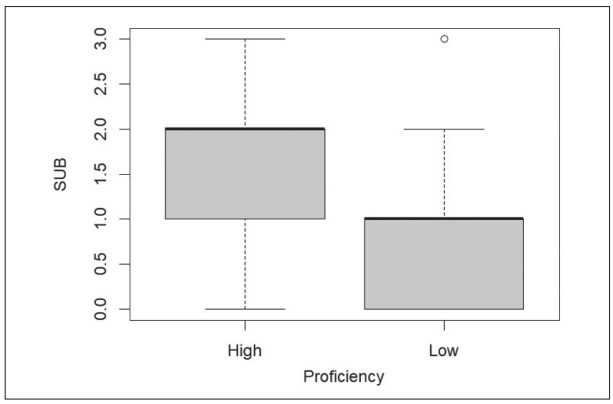
■表8: 非テキスト情報の出現率

	High Proficiency			Low Proficiency		
	n	*M*	*SD*	*n*	*M*	*SD*
SUB	17	1.59	0.91	19	0.74	0.78
ADD	17	1.65	0.97	19	1.00	1.08
SUM	17	0.94	0.80	19	0.47	0.68

Note. SUB: Substitution, ADD: Addition, SUM: Summary

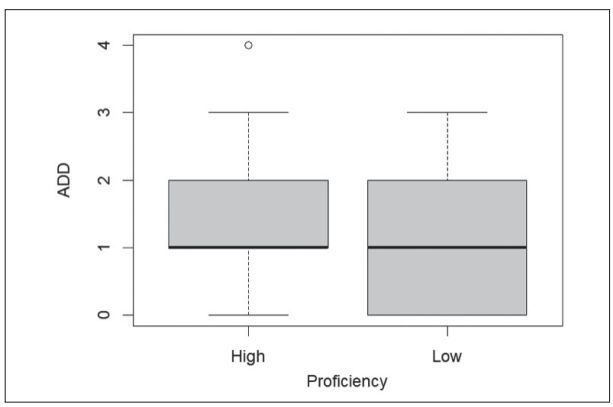
コルモゴロフ・スミルノフ検定の結果，SUBについては，どちらの熟達度群でも有意な正規性を満たしており，高熟達度群は$D(17) = .205$, $p = .473$, 低熟達度群は$D(19) = .267$, $p = .134$であった。ADDについても，どちらの熟達度群でも有意な正規性を満たしており，高熟達度群は$D(17) = .271$, $p = .163$, 低熟達度群は$D(19) = .291$, $p = .08$であった。SUMについては，高熟達度群は有意な正規性を満たしており[$D(17) = .225$, $p = .353$]，低熟達度群は有意な正規性を満たしていなかった[$D(19) = .383$, $p = .008$]。よって，SUBとADDに関してはパラメトリック検定，SUMについてはノンパラメトリック検定を採用した。

熟達度を考慮し，再話プロトコルにおける非テキスト情報の出現率を比較するため，各カテゴリーについて，対応なしt検定（SUB・ADD）及びウィルコクソンの順位和検定（SUM）を行った。t検定を使い，SUBの出現率を高熟達度群/低熟達度群で比較した結果，$t(31) = 2.90$, $p < .01$, $d = .977$ [95%CI = .26, 1.70]で，高熟達度群の方が有意にSUBの出現率が高かった。



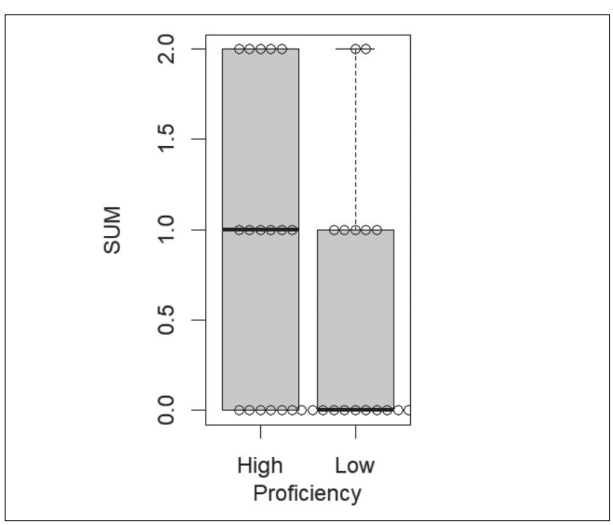
■図6：SUBの出現率（箱ひげ図）

ADDについても同様に*t*検定を行った結果，$t(34) = 1.85$, $p = .073$, $d = .613$ [95%CI = -.081, 1.31]で，2つの熟達度間で有意な差は見られなかった。



■図7：ADDの出現率（箱ひげ図）

SUMについて，ウィルコクソンの順位和検定を用いて熟達度間で差があるかを検証した結果，2つの熟達度間で差はなかった（$W = 213$, $p = .085$, $r = .172$）。



■図8：SUMの熟達度別蜂群図

次に，各タスク条件ごとの非テキスト情報の出現率を計算した。結果を表9〜11に示す。

■表9: SUBの出現率

	n	*M*	*SD*
アンダーライン	13	1.15	1.03
キーワード	12	1.25	0.83
図式化	11	1.00	0.95

■表10: ADDの出現率

	n	*M*	*SD*
アンダーライン	13	1.15	1.10
キーワード	12	1.42	0.95
図式化	11	1.36	1.15

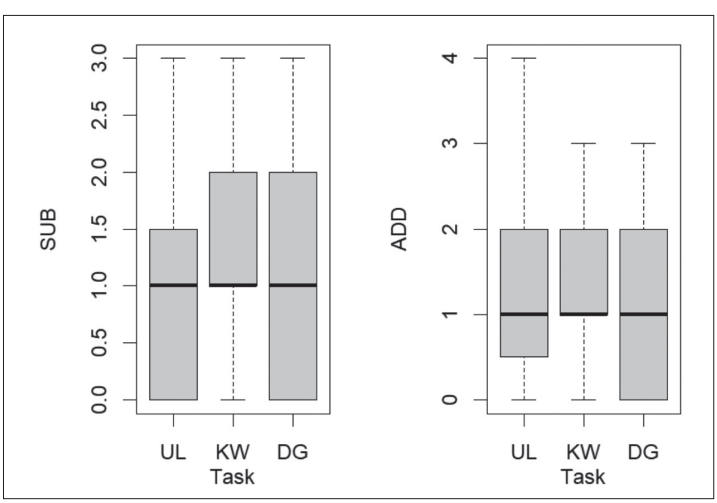
■表11: SUMの出現率

	n	*M*	*SD*
アンダーライン	13	0.62	0.74
キーワード	12	0.83	0.80
図式化	11	0.64	0.77

コルモゴロフ・スミルノフ検定の結果，SUBについては，有意な正規性を満たしており，$D(36) = .223$, $p = .055$であった。ADDも有意な正規性を満たしており，$D(36) = .229$, $p = .045$であった。SUMについては有意な正規性を満たしておらず，$D(36) = .343$, $p < .01$であった。よって，SUBとADDに関してはパラメトリック検定，SUMについてはノンパラメトリック検定を採用した。

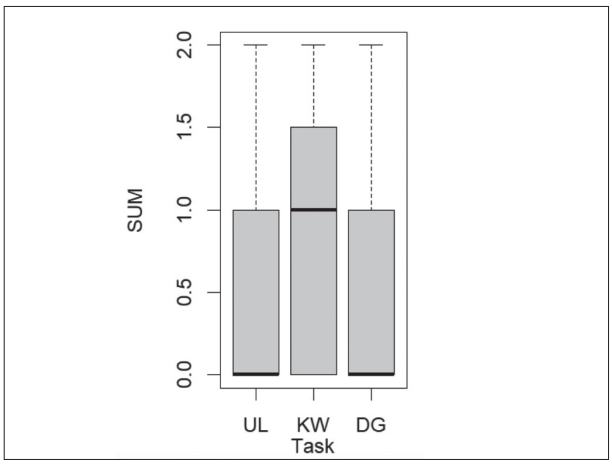
各タスク条件におけるSUBの出現率について，対応なしの一元配置分散分析を行った。その結果，$F(2, 33) = .189$, $p = .830$, $\eta_G^2 = .011$であり，タスク間に有意差は見られなかった。

同様に，ADDについても対応なしの一元配置分散分析を行った結果，$F(2, 33) = .194$, $p = .824$, $\eta_G^2 = .012$であり，タスク間に有意差は見られなかった。



■図9: SUB, ADDの出現率(タスク別)

各タスク条件におけるSUMの出現率について，クラスカル・ウォリスの順位和検定を行った。その結果，$X^2(2) = .565$, $p = .754$で，タスク間に有意な差は見られなかった。



■図10：SUMの出現率（タスク別）

3.4 アンケート

テキストの難易度と実験で行ったノートテイキングの効果を調べるために，4つの項目からなるアンケートを実施した。Q1とQ2は，英文の内容を知っているかどうか，英文の難易度に関する質問である。Q3は，「指示されたタスクは読解に役立ったか」，「Q3の選択理由について，なぜそう思うのか」であった。結果を表12, 13に示す。

■表12:

	Scale 1		*Scale 2*		*Scale 3*		*Scale 4*		*Scale 5*	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Q1	23	63.9	6	16.7	1	2.8	6	16.7	0	0
Q2	1	2.8	9	25	8	22.2	13	36.1	5	13.9
Q3	1	2.8	5	13.9	4	11.1	21	58.3	5	13.9

Note. Q1. この英文の内容を知っていましたか（1: 知らなかった〜 5: 知っていた） Q2. テキストの難易度について（1: 簡単だった〜 5: 難しかった）Q3. 読解前に指示されたタスク（アンダーライン / キーワード抽出 / 図式化）は，テキストを理解するために役に立ったと思いますか（1: 役に立たなかった〜 5: 役に立った）

■表13:

	High	Low
	M(*SD*)	*M*(*SD*)
Q1	1.88 (1.23)	1.58 (0.99)
Q2	3.59 (1.03)	3.11 (1.07)
Q3	3.88 (1.02)	3.47 (0.88)

Q3「読解前に指示されたタスク（アンダーライン / キーワード抽出 / 図式化）は，テキストを理解するために役に立ったと思いますか」について，各タスク条件ごとの回答を表14に，各スケールの平均及び標準偏差とスコアを表15に示す。

■表14: Q3の回答(タスク別)

	Scale 1		*Scale 2*		*Scale 3*		*Scale 4*		*Scale 5*	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
UL	0	0	1	7.7	2	15.4	7	53.8	3	23.1
KW	1	8.3	3	25	1	8.3	6	50	1	8.3
DG	0	0	1	9.1	1	9.1	8	72.7	1	9.1

■表15: Q3の回答の平均値(標準偏差)

	M(*SD*)
UL	3.92 (0.83)
KW	3.25 (1.16)
DG	3.82 (0.72)

アンケートの結果，各タスク条件では，アンダーライン条件が最も役に立っていると感じている人が多く，キーワード条件が最も少なかった。

Q4の自由記述の質問では，Q3の回答理由について尋ねた。自由回答の結果を表16〜18にまとめる。アンダーライン条件で役に合った / やや役に立ったと回答した協力者からは，アンダーラインのタスク指示を受けることで，テキストの重要な箇所を意識しながら読むことができた，集中して読めた，自分でアンダーラインを引いた箇所は，他の文よりも記憶に残りやすくなったなどの意見が挙げられた。3条件の中でもっとも役に立ったと感じた人が少なったキーワード条件では，役に立ったと回答した協力者からは，キーワードの抽出を行う事で，内容の整理や理解につながった，要点を意識して読めたなどの意見が挙げられた。一方で，役に立たなかったと回答した協力者からは，メモに記憶を任せていて何も覚えていなかった，単語そのものに文脈の要素が入っていないため，記述の役に立たなかったなどの意見が挙げられた。

■表16: アンダーライン条件

Scale	理由
5	【上位群】 • 初見の単語でも，使われている複数箇所をマークすると，文脈から意味をある程度推測できたほか，キーワードを抽出することで文章の構造を探ることができたため • テキストの重要な部分を意識し，順序だてて理解できたから 【下位群】 • 線を引いてくださいと言われたことで，集中して，どこが大事そうかに対してタスクできたから
4	【上位群】 • キーワードが目立つようになり，より脳に残りやすくなったから(覚えられるようになった) • テキストの重要な部分はどこなのかを考えながら読むことを意識したから • アンダーラインを自分で引くことで，他の文章よりも記憶に残りやすいから 【下位群】 • テキストの重要な部分について整理できたから，アンダーラインを引くことで記憶に残りやすくなるから • アンダーラインを引いたところが全てではないが記憶に残っていたから • メインのお題に関連している内容と，関連性の低いものに分けて考えることができた

■表16: アンダーライン条件

Scale	理由
3	【下位群】 • テキストの重要と思われる部分について整理することはできたが，実際に問題を解くときにその箇所を見る事が出来ず，あまり思い出せなかったので役に立ったのかどうかはっきりしないため • 特に必要性を感じなかったため
2	【下位群】 • 文章の量や改行などによって話のまとまりをつかみやすかったため，アンダーラインはあまり必要ないと思ったから

■表17: キーワード条件

Scale	理由
5	【上位群】 • 一度自分の手で書くことによって，内容を確認したり思い出したりすることができたから
4	【上位群】 • 数字や固有名詞を書き留めることで，流れを思い出しやすくすることができたと感じたから • それぞれの内容ごとに整理できたから • タスクをこなすように指示されることによって，集中力が上がるから • 自分で要約文を書くときに頭にキーワードが何個か残っていたから 【下位群】 • キーワードを書き出して可視化できる状態で整理することで，頭の中を整理することができたため • キーワードの抽出を行う事で，内容の整理や理解につながったから • 要点を意識しながら読めたから
3	【上位群】 • キーワードを抽出したことで文章の内容について理解を進めることができた。一方で詳細な説明事項についてはメモを取ったため，文章全体を要約する上では細かな内容を折り込むことはできなかった
2	【上位群】 • キーワードの抽出のみだと，未知語の推測が文脈に沿ってできないため，文章を後から振り返れないときに，文章理解の時に日本語で書き出すことができない 【下位群】 • 固有名詞が全く頭に入っていなかった。メモを見たら文章を書くのは簡単だと思っていたが，メモに記憶を任せていて何も覚えていなかった
1	【上位群】 • 単語そのものに文脈の要素が入っていないため，記述の役に立たなかった

■表18: 図式化条件

Scale	理由
4	【上位群】 • 英語を読んでいるときに記憶を保持するのが難しいので，図式化によりその前に読んだ中身を忘れないようにできたから • 段落ごとの関係性についての理解が深まったから • 話の要点を図式化することで，どこで理解がつまずいているのかが分かりやすくなったから • テキストの内容を整理できたため • 各パラグラフごとの内容を簡潔にノートにとることで，全体の流れを把握することができたから 【下位群】 • 内容を整理しやすくなった • テキストの大事な部分についての記憶定着が起きていたと思うため • 順序的に内容を理解することができたから • テキストの因果関係を目で見て理解することができたから
3	【下位群】 • 指示内容は意識せずに文章の内容をまとめたから
2	【下位群】 • 内容の整理はできたと思うが頭には入っていなかった

3.5 読解中タスクのメモや再話プロトコルの質的な分析

3.5.1 アンダーライン

アンダーライン条件の協力者が引いたアンダーラインの数を計算したところ，上位群の平均が13.4（n = 5），下位群の平均が5.5（n = 6）と，上位群の方が多くのアンダーラインを引いていることがわかった。また，アンダーラインを引いた箇所を観察したところ，再生率が高い協力者は単語や句単位でラインを引き，再生率の低い協力者は節単位の長いアンダーラインを引いている傾向が明らかになった。

■表19: 協力者のアンダーライン

再生率：高	再生率：低
Every morning and evening, Howler monkeys have a group howl, a noise that can be heard three miles away. This loud howl keeps other animals away from the troops' food source.	Every morning and evening, Howler monkeys have a group howl, a noise that can be heard three miles away. This loud howl keeps other animals away from the troops' food source.
The Saddle Bronc rider must stay on the horse at least 8 seconds. The rider is disqualified if he touches any part of the horse or his own body with his free hand.	The Saddle Bronc rider must stay on the horse at least 8 seconds. The rider is disqualified if he touches any part of the horse or his own body with his free hand.

また，再生率が高い協力者と低い協力者の再話プロトコルを比較したところ，再生率が高い協力者はアンダーラインと再話プロトコルで一致している箇所が多いのに対して，再生率が低い協力者はアンダーラインを引いた箇所を再話に反映できていないことがわかった。

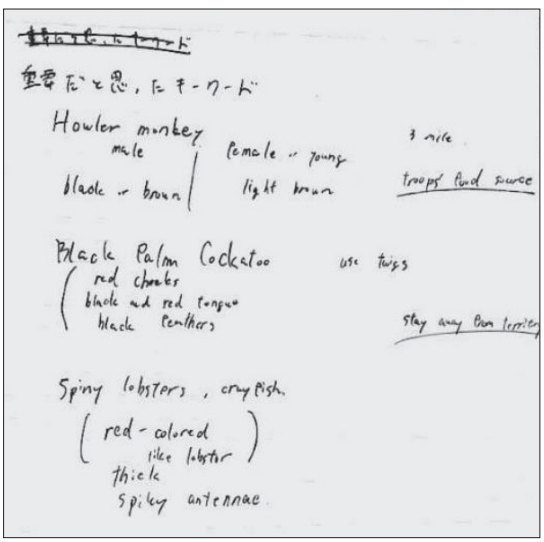
■表20: 再生率が特に高い/低い協力者4名のアンダーラインと再話の一致数

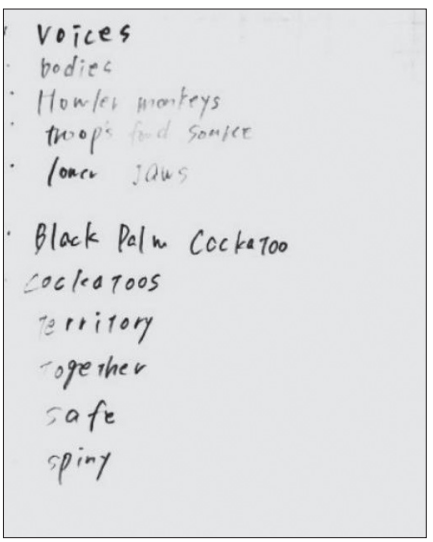
再生率:高い			再生率:低い		
n	IU	非テキスト情報	*n*	IU	非テキスト情報
11	5	2	3	0	0
17	9	2	6	0	0

※*n* = アンダーラインの数

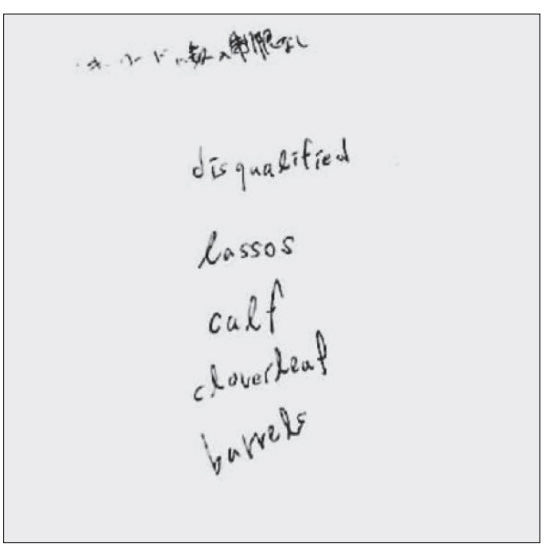
3.5.2 キーワード条件

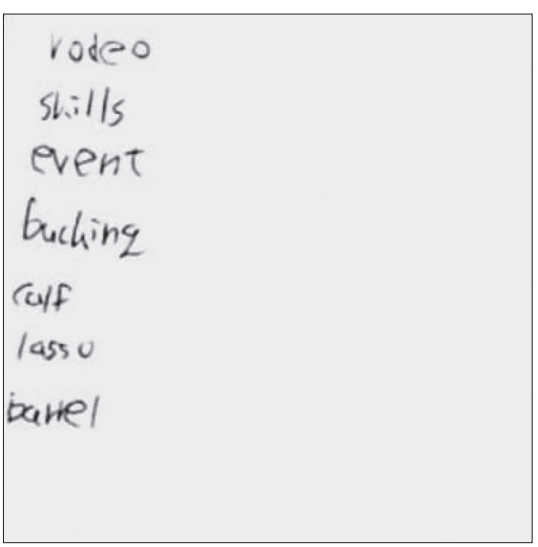
キーワード条件の協力者のメモを観察したところ，①テキストの意味のまとまりに合わせて，抽出したキーワードの配置を工夫したもの，②単に重要なキーワードを抜き出してメモしていくもの，という大きく2つの特徴が見られた。熟達度によらず，再生率が低い協力者は②の形式でメモしている傾向があった。





■図11:再生率が高かった協力者のメモの例





■図12:再生率が低かった協力者のメモの例

また，再生率が高い協力者と低い協力者の再話プロトコルを比較したところ，再生率が高い協力者はキーワードを基にIUと同じまたは言い換え表現を使って再生できているのに対し，再生率が低い協力者はキーワード抽出した部分は再話プロトコルに出現しているものの，IUの2/3以上の情報を再生できず，得点にはなっていないケースが多く見られた。

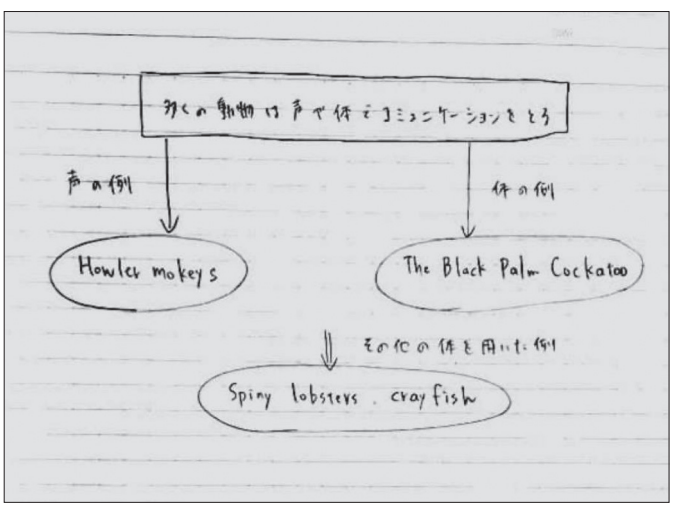
■表21: 再生率が特に高い/低い協力者4名のキーワードと再話の一致数

再生率：高い			再生率：低い		
n	IU	非テキスト情報	*n*	IU	非テキスト情報
12	4	3	5	0	0
5	4	0	21	0	0

Note. *n* = アンダーラインの数

3.5.3 図式化条件

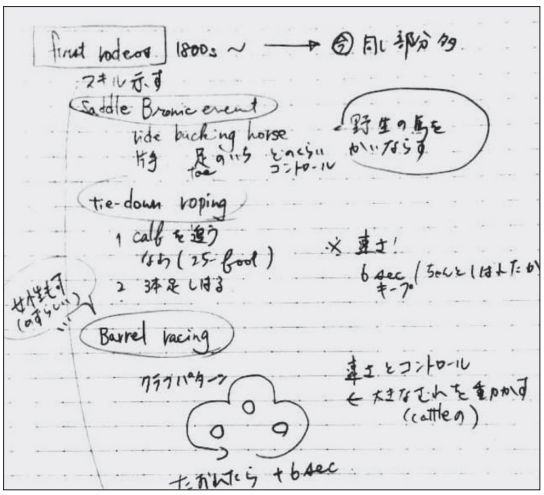
図式化条件の協力者は，全体的に再生率が高い傾向にあった。メモを観察したところ，日本語/英語，矢印の方向など，図式化の方法には個人差があるものの，本文の大枠の構成ごとに内容をまとめている傾向があった。例えば，テキストAは動物のコミュニケーションについて，導入→Howler monkeysの例→Black Palm Cockatooの例→ロブスターやザリガニの例という構成になっているが，図13のように，矢印を使って簡潔にまとめている。図13のメモを作成した協力者は熟達度が下位群であったが，再話では高い再生率を得ていた。



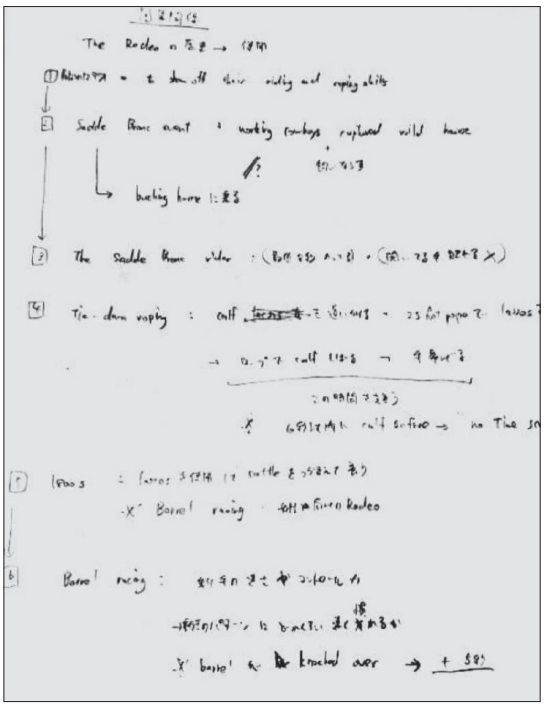
■図13：図式化条件の協力者のメモの例

図14のように，大枠の構成に詳細情報も加えた図式化も多く見られた。テキストBはロデオの3つの種目と採点方法について，導入→Saddle Bronc→Tie-down roping→Barrel racingの順で説明されているが，種目の過程について番号で示す，詳細情報を吹き出しで説明する，Barrel racingのクローバーリーフパターンを絵でまとめるなどの工夫が見られた。また，同協力者の再話を観察したところ，3つの種目を順序だてて説明するために，「ここでは，3つの内容が例として紹介されていた。」と本文には文字通り書かれていない情報を補ってまとめていた。

一方で，図15からわかるように，再生率が低かった協力者は，情報量が多く，すべてのパラグラフを要約する形で図式化する傾向にあった。



■図14：図式化条件の協力者のメモの例



■図15：再生率が低かった協力者のメモ

4 考察

4.1 全体的な考察

本研究では，読解中の活動であるノートテイキングがテキスト内容の記憶にどのような影響を与えるのかを検証することを目的に，アンダーライン/ キーワード抽出/図式化のいずれかの読解タスクを与え，読解後の再話プロトコルを比較した。二元配置分散分析の結果，熟達度上位群は有意に下位群よりも再話の成績が高かったが，タスク条件との交互作用は得られなかった。タスク別の再話成績では，平均値は図式化条件＞キーワード条件＞アンダーライン条件の順に高く，図式化条件とアンダーライン条件の間には有意な差が得られた。また，非テキスト情報について，熟達度上位群の方が有意にSUBの出現率が高かったが，ADDやSUMについては熟達度間で差はなく，各タスク条件間でも有意な差は見られなかった。以上の結果を踏まえ，各RQについて考察を行う。

RQ1 ノートテイキングの種類は再話成績に影響を与えるか

最初の研究課題は，ノートテイキングの種類が再話成績に有意な影響を与えるかどうかを理解することを目的とした。二元配置分散分析の結果，ノートテイキングの種類の主効果は統計的に有意ではなかった。この結果は，本研究の範囲内では，異なる種類のノートテイキング方法（アンダーライン，キーワード抽出，図式化）は，再話成績に統計的に有意な差をもたらさなかったことを意味している。この結果は筆者の予想と反していた。メモの取り方を変えれば，記憶の定着に影響を与え，その結果，再話成績にも影響を与えるのではないかと考えられるからである。しかし，有意な効果が見られなかったことから，メモの取り方とリテリングの成績の関係は，想定されていたよりも微妙なものであることが示唆された。

この有意性の欠如にはいくつかの要因が考えられる。第一に，この研究で用いられたタスクが特殊であったため，各ノートテイキング方法の違いを効果的に引き出すことができなかった可能性がある。例えば，図式化はしばしば視覚的表現と要約に関連しているが，協力者は様々な形式で図式化を行っていたため，本研究の図式化タスクが，この手法の可能性を完全に捉えていなかった可能性がある。アンダーラインやキーワード抽出のタスクについても同様である。このように，タスクそのものが，それぞれのノートテイキング手法の潜在的な利点を明らかにするために最適化されていなかった可能性があるため，今後の研究では，指示をより細分化し，各タスク条件を比較する必要がある。

さらに，学習スタイルや嗜好の個人差が結果に影響する可能性もある。例えば， 視覚的に理解することを好む学習者は図式化を好み、言語的な理解を好む学習者はキーワードや重要なフレーズをメモすることを好むことが考えられ、こういった個人差が再話におけるノートテイキングの一般的な効果に影響を与える可能性は、要因として考慮する必要がある。

内容理解タスクのタイミングを考慮することも重要である。ノートテイキングを行った後すぐに再話するのと，遅れて再話するのとでは，結果が異なる可能性がある。即時再話はノートの鮮明さや新鮮さに影響されるかもしれないが，遅延再話は異なるノートテイキング方略が長期記憶に与える影響をテストすることができる。本研究では，即時再話に焦点を当てたため，研究結果の範囲が限定され，特定のノートテイキング方法の長期にわたる潜在的な利点を十分に把握できない可能性がある。

結論として，最初の分析では，ノートの取り方が再話の成績に及ぼす有意な影響は明らかにならなかったが，これは特定のノートの取り方の潜在的な利点を必ずしも否定するものではない。この結果から，メモの取り方と記憶想起の関係は，課題デザイン，個人差，評価のタイミングなど様々な要因に影響されることが示唆される。

RQ2　熟達度はRQ1の結果に影響を与えるか

第二の研究課題は，協力者の熟達度がノートテイキングの種類と再話成績との関係に影響を与えるかどうかを探ることであった。研究の結果，熟達度が再話パフォーマンスに有意な主効果を及ぼすことが明らかになった。熟達度の高い協力者は，熟達度の低い協力者に比べて，優れた再話成績を示した。この結果は，テキストから情報を理解し，効果的に想起する上で，熟達度が極めて重要な役割を果たすことを強調している。

読解熟達度が再話の成績に大きな影響を与えることは，第二言語学習の分野で確立された研究と一致している。熟達度が高い人ほど，ワーキングメモリ運用能力が高く（Geva & Ryan, 1993），語の意味や文法解釈などの低次の処理が自動化されている。よって，より高次の処理に認知資源を当てることができ，複雑なテキスト情報を処理し，保持することができる（門田・野呂，2001）。さらに，熟達度の高い学習者はメタ認知方略が優れている傾向があり，理解プロセスをより効果的に調整することができる（Block, 1992）。したがって，本研究の結果は，熟達度が読解とその後の記憶想起を成功させるための基礎的要素として機能するという考え方を補強するものである。

RQ3　どのようなノートの取り方が再話を促進させるか

第三の研究課題は，どのようなノートテイキングがより効果的に再話を促進するかを明らかにすることである。熟練度とノートテイキングのタスクのタイプの間の交互作用は，有意な効果をもたらさなかった。しかし，異なるタスクにおける協力者のノートテイキング行動をより深く検討することで，ノートの取り方が再話の成績に影響を与える可能性が示唆され，貴重な洞察が得られた。詳細については，4.2節で述べる。

4.2 各タスク条件でのメモの違い

アンダーライン条件では，熟達度上位群の方がより多くのアンダーラインを引いており，再生率が高い協力者は単語や句レベルで印をつけていることがわかった。一方で，再生率が低い協力者は，節単位の長いアンダーラインを引く傾向にあり，重要な情報を特定することに苦労している可能性が示唆された。また，アンダーラインを引いた箇所と再話プロトコルを照らし合わせると，再生率が低い協力者はアンダーラインを引いた箇所を再話プロトコルに含めておらず，再生率が高かった協力者は，アンダーラインを引いた箇所を再話プロトコルに多く含める傾向にあった。このことから，教育現場においてノートテイキングを指導する際は，熟達度の低い人が重要な要素に集中できるよう，より足場となるアプローチが必要であることが示唆された。

キーワード条件の協力者のノートを観察したところ，再生率が高い協力者はテキストの意味のまとまりに合わせてメモの位置を工夫していたのに対して，再生率が低い協力者は，キーワードを横または縦にそのまま抽出している傾向にあった。先行研究では，情報を書き写すことが浅い処理を伴うのに対し，情報を整理することは深い意味的処理を伴い，長期的な記憶を促進する（e.g., Bui et al, 2012）と考えられており，同じキーワード抽出というタスクの中でも，書き写しただけの場合よりも，重要なキーワードを整理してメモする方が，文章の流れを整理しやすくなった可能性が考えられる。また，アンケートでは，肯定的な意見としては「キーワードの抽出を行う事で，内容の整理や理解につながった」「要点を意識しながら読めた」などが挙げられたが，「単語そのものに文脈の要素が入っていないため，記述の役に立たなかった」などの意見も挙げられた。このことから，頭の中，もしくはメモをとりながら，文脈も含めて情報を整理できた協力者はテキストの主要なアイデアを抽出し，キーワードを活用できた可能性があるが，単に重要だと思うキーワードを抽出しただけの場合は，核となる概念の特定や文構造の理解に困難を抱えていた可能性があることが示唆された。

図式化条件の協力者のノートを観察したところ，再生率が高い協力者ほど，英文の大枠の構造に詳細情報を付け足して図式化しており，再生率が低かった協力者は，すべてのパラグラフの情報を要約する形で，多くの情報を図式化している傾向にあった。熟達度が下位群の協力者でも，主要な構造をメモできていた協力者は再生率が高かった。このことから，図式化する際には，多くの情報をメモすることがより高い再生率につながるわけではないことが示唆された。先行研究では，読んだことをすべて書き写す読み手のワーキングメモリはアクティブであるため，入ってきた知識を分析することができない（Hill & Miller, 2006）ことが明らかになっている。このことから，図式化においても，逐語的なメモをたくさん残すことで，高次的な処理や意味を保持するための容量が残されていなかった可能性がある。アンケートでは，肯定的な意見として「話の要点を図式化することで，どこで理解がつまずいているのかが分かりやすくなった」「段落ごとの関係性についての理解が深まった」などの肯定的な意見が挙げられた。先行研究では，ノートテイキングとメタ認知には関係があることがわかっており，読解中にノートを取る学習者は，ノートを取らない学習者よりも，学習内容の理解度をより正確に把握し，学習課題を調整できることが明らかになっている（Brozo et al., 1996）。このことから，図式化を行い，テキスト内容を整理する過程で，自分の理解度を把握できたことで，理解度に応じた読みを行い，結果的に再話でよい成績を収めることができた可能性が示唆された。一方で，「内容の整理はできたと思うが頭には入っていなかった」という意見も挙げられた。認知負荷の観点から，先行研究では協力者自身が読解中に図を作成することは効果がないと結論づけた研究も見られる（e.g., Stull & Mayer, 2007）。これらの先行研究を踏まえると，図式化を行う際，テキストベースの理解や図の作成に認知容量を使用してしまい，記憶として保持することが困難となってしまった可能性がある。以上の結果から，教育現場で指導する際は，生徒の個人差に留意し，図式化のフレームワークを提示するなどの介入を行う必要があることが示唆された。

キーワード条件，図式化条件で比較的再話の再生率が高かった要因として，メモを残す際にテキストの構造を理解できていたかが重要であったと考えられる。これは第二言語学習者を対象とし，図式化が

再話成績にどのような影響を及ぼすのか調査した高橋(2017)の結果と一致している。先行研究では，アンダーラインは項目間の関連性よりも個々の項目に注意が向けられる(Dunlosky et al., 2013)ことが示唆されており，テキストの表面的な処理や強調した部分の暗記はサポートされるが(Mayer, 1996)，文章全体の構造や関係性については理解が促進されなかった可能性がある。今回は内容理解を測定するタスクとして再話を採用したため，文章構造を理解し，まとめる力が必要とされ，アンダーラインが理解促進につながらなかった可能性があるが，重要だと思う箇所にアンダーラインを引くだけでなく，英文同士の関係を示す矢印を記入することも認めた場合，よりアンダーラインの効果が得られる可能性がある。

4.3 限界点

本研究には，いくつかの限界点があった。まず，タスク条件が3つあることを考慮すると，サンプル数が36と少なく，有意差が生じにくくなってしまった可能性がある。また，協力者が1つの条件で1つのテキストを読む協力者間要因を採用したことにより，ノートテイキングをしなかった場合の再話成績は測定できなかったため，本研究からは，各協力者がノートをとった場合に，ノートをとらなかった場合と比較してどれだけ記憶が促進されたかはわからなかった。今後の研究では，ノートの取り方やタスクの種類を考慮し，ノートをとらなかった場合と比較してどれほどの違いがあるのか調査する必要がある。また，読解後の理解タスクの種類(多肢選択式/再話など)によって，効果的なノートの取り方やノートテイキングの種類が異なる可能性があることから，様々な要因を考慮した実験デザインにより，効果的なノートテイキングを検討する必要がある。

4.4 教育的示唆と今後の展望

本研究の結果は，研究と教育実践の双方に示唆を与えるものである。第一に，熟達度が有意に影響したことから，読解やノートの取り方に関する介入を設計する際に，個人差を考慮する必要性があることが示唆された。熟達度に応じて指導方略を調整することで，より効果的な学習成果が得られる可能性がある。

再話の成績にタスクの種類が有意に影響しなかったことから，本研究からは，ノートテイキングの種類の選択は再話成績の決定要因ではない可能性が示唆されたが，今後の研究では，ノートテイキングの種類と様々な認知的課題との間のより具体的な相互作用を探る必要がある。例えば，本研究では読解後の課題として再話を採用し，学習者は日本語を用いた筆記形式で課題に取り組んだが，口頭で，英語を用いて再話を行った場合には，つなぎ言葉や正確性の観点において，ノートテイキングの種類による影響が観察される可能性がある。さらに，個人の学習スタイルや嗜好の影響を調査することで，様々な学習者がノートテイキング方略にどのように取り組み，そこからどのような利益を得ているかについて，より豊かな理解を得ることが期待される。

本研究では，ノートテイキングの質的観察から，再話成績の異なる個人がどのようにノートテイキングに取り組んでいるかについての洞察が得られた。この知識は，学習者が効果的にノートテイキング技術を活用できるように導く，的を絞った指導方略の開発に役立てることができる。例えば，事前にテンプレートを提供し，構造を整理しやすくするような指導を行うことで，熟達度の低い学習者がノートテイキングにより深く取り組み，理解力や記憶力の向上につながる可能性がある。

5 結論

本研究は，ノートの取り方，熟達度，再話成績の間の複雑な関係についての理解を深めるものである。実験では，読解中の活動であるノートテイキングがテキスト内容の記憶にどのような影響を与えるのか

を検証することを目的に，アンダーライン/ キーワード抽出/ 図式化のいずれかの読解タスクを与え，読解後の再話プロトコルを比較した。二元配置分散分析の結果，各タスク条件で再話の成績に有意性は見られなかったが，熟達度の主効果は有意であることが明らかになった。また，協力者のノートと再話プロトコルの質的な分析から，ノートの取り方と再話成績には関係がある可能性があり，生徒にノートテイキングを指導する際には，英文の構造がわかるようなノートをとることを伝えることで，より強固な心的表象を構築できることが示唆された。これらの結果は，認知プロセスを調査する際に個人差を考慮することの重要性を強調し，教育者が方略的なノートテイキングの実践を通じて読解力と記憶保持力を高めるための実践的な洞察を提供するものである。

謝辞

この論文の完成にあたり，多くの方々からのご支援とご協力を賜りましたことを心より感謝申し上げます。以下に謝辞を捧げさせていただきます。まず，研究実施にあたり，公益財団法人 日本英語検定協会および関係者の皆様，選考委員の先生方に心より感謝申し上げます。貴重なご支援と協力により，実施するにあたり必要な環境を整えていただきました。次に，実験実施および報告書の指導に関して，助言者である寺内一先生に深く感謝申し上げます。ご指導とアドバイスにより，研究の進行および論文の執筆において大変な助けとなりました。また，研究を進める上で貴重な助言を下さった筑波大学の卯城祐司先生をはじめ，分析のアドバイスや採点のお手伝いをしてくださった研究室の先輩・同期の皆様にも，心からの感謝を申し上げます。最後に，実験に協力していただいた実験協力者の皆様にも，心より感謝申し上げます。以上の皆様のご支援とご協力のおかげで，この研究を遂行し，論文を完成させることができました。改めて，心より感謝申し上げます。

引用文献

Bahrami, F., & Nosratzadeh, H. (2017). The effectiveness of note-taking on reading comprehension of Iranian EFL learners. *International Journal of Applied Linguistics and English Literature, 6*(7), 308-317.

Block, E. L. (1992). See how they read: Comprehension monitoring of L1 and L2 readers. *TESOL Quarterly, 26,* 319-343

Boyle, J. R. (2010). Strategic note-taking for middle-school students with learning disabilities in science classes. *Learning Disability Quarterly, 33,* 93–109.

Brozo, W. G. & Stahl, N.A., & King, J.R. (1996). Tracking Thinking While Reading: A Strategy for Improving Metacomprehension. *Research and Teaching in Developmental Education, 13*(1), 19-31.

Bui, D. C., Myerson, J., & Hale, S. (2012). Note-Taking With Computers: Exploring Alternative Strategies for Improved Recall. *Journal of Educational Psychology.* Advance online publication.

Chang, W. C., & Ku, Y. M. (2015). The effects of note-taking skills instruction on elementary students' reading. *The Journal of Educational Research, 108*(4), 278-291.

Cook, L. K., & Mayer, R. E. (1983). Reading strategies training for meaningful learning from prose. In *Cognitive strategy research: Educational applications.* 87-131.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63,* 8–14. doi:10.1037/h0032243

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public interest, 14*(1), 4-58.

Hedgcock, J., & Ferris, D. (2009). *Teaching readers of English: Students, texts, and contexts.* New York, NY: Routl.

Geva, E., & E. B. Ryan. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning, 43,* 5-42.

Hill, J. D., & Miller, K. B. (2006). Classroom instruction that works with English language learners. Virginia, USA: Association for Supervision and Curriculum Development.

平井明代・岡秀亮・草薙邦広（編著）(2022).『教育・心理研究のためのRによるデータ分析―論文作成への理論と実践集』東京：東京図書.

Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language, 25*(2), 98-125.

市澤慧太郎, 吉野巌. (2018). グラフィック・オーガナイザーが日本人英語学習者の読解方略に与える効果. 北海道教育大学紀要. 教育科学編.

Ikeno, O. (1996). The effects of text-structure-guiding questions on comprehension of texts with varying linguistic difficulties.

引用文献

JACET Bulletin, 27, 51-68.

門田修平, 野呂忠司. (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.

Kiewra, K. A., Benton, S. L., & Lewis, L. B. (1987). Qualitative aspects of notetaking and their relationship with information-processing ability and academic achievement. *Journal of Instructional Psychology, 14,* 110- 117.

Kimura, Y. (2022). Text processing and memory in EFL reading: The role of relevance instructions. *Reading in a Foreign Language, 34*(1), 41.

小林敬一. (2000). 共同作成の場におけるノートテイキング・ノート見直し. *教育心理学研究* , 48(2), 154-164.

Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30*(2), 242-262.

久保田章・林伸昭(編著)(2019).『授業力アップのための英語教育学の基礎知識』東京:開拓社.

Kucer, S.B. & Silva, C. (1999). The English literacy development of bilingual students within a transition whole-language curriculum. *Bilingual Research Journal, 23,* 339-364.

Kucer, S.B. (2009) *Dimensions of Literacy: A Conceptual Base for Teaching Reading and Writing in School Settings,* 3rd edn. New York: Taylor and Francis/Routledge.

Kucer, S. B. (2011). Going beyond the author: What retellings tell us about comprehending narrative and expository texts. Literacy, 45(2), 62-69.

Lonka, K., Lindblom-Ylänne, S., & Maury, S. (1994). The effect of study strategies on learning from text. *Learning and Instruction, 4*(3), 253-271.

Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educational Psychology Review, 8*(4), 357-371.

Mežek, Špela. (2013). Learning terminology from reading texts in English: The effects of note-taking strategies. *Nordic Journal of EnglishStudies, 13*(1), 133-161.

Nanna, I., Louise, B., & Claus, B. (2021). Interacting with academic readings — A comparison of paper and laptop, *Social Sciences & Humanities Open, 4*(1), 2590-2911.

Özçakmak, H. (2019). Impact of Note Taking during Reading and during Listening on Comprehension. *Educational Research and Reviews, 14*(16), 580-589.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New Yolk: Oxford University Press.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Puji, S. (2019). Note Taking Strategy: Its effect on Reading comprehension, *Journal of English Language Studies, 4*(1), 32-40.

Rahmani, M., & Sadeghi, K. (2011). Effects of note-taking training on reading comprehension and recall. *Reading, 11*(2), 116-128.

Robinson, D. H., & Molina, E. (2002). The relative involvement of visual and auditory working memory when studying adjunct displays. *Contemp. Educ. Psychol. 27*(1), 118-131.

齋藤ひとみ, 源田雅裕. (2008). ノートテイキングにおける方略使用の効果に関する検討. *日本教育工学会論文誌*, 31(Suppl.), 197-200.

関友作. (1997). テキストの内容把握に対する箇条書とキーワード強調の影響. 日本教育工学雑誌. *日本教育工学会*, 21, 17-20.

Slotte, V., & Lonka, K. (1999). Review and process effects of spontaneous note-taking on text comprehension. *Contemporary Educational Psychology, 24,* 1-20.

Stull, A. T., & Mayer, R. (2007). Learning by doing versus learning by viewing: Three Experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology, 99*(4), 808-820. https://psycnet.apa.org/doi/10.1037/0022-0663.99.4.808

Suzuki, A. (2007). The Effects of simultaneous display of information by a graphic organizer in EFL reading. *JACET Journal, 45,* 47-61

高橋亜紀子. (2017). 中・上級日本語学習者は文章をどのように読んでいるのか: 図の作成・再話タスクの分析を通して. *宮城教育大学紀要*, 51, 229-240.

Tilstra, J, McMaster, K. (2013). Cognitive processes of middle grade readers when reading expository text with an assigned goal. *Learning and Individual Differences, 28,* 66-74.

卯城祐司(編)(2009)『英語リーディングの科学ー「読めたつもり」の謎を解く』東京:研究社.

魚崎祐子, 伊藤秀子, 野嶋栄一郎. (2003). テキストへの下線ひき行為が内容把握に及ぼす影響. *日本教育工学雑誌*, 26(4), 349-359.

Williams, Robert L. and Alan C. Eggert. (2002). Notetaking in College Classes: Student Patterns and Instructional Strategies. *The Journal of General Education, 51*(3), 173-19.

資料：熟達度テスト[1]

An Extraordinary Machine

Most of the machines that people in developed nations use were invented during the last 200 years. They make tasks easier for people and give them more time for other tasks and for leisure. However, which of these machines has changed society the most? Even though people spend more time with their TVs, computers, and smartphones, some historians argue that the impact of these inventions has been small compared with that of washing machines.

Before washing machines, clothes and sheets were washed by hand. For most of history, this has involved carrying the laundry to a river or a lake, wetting it, and rubbing it with rocks, sand, or soap to remove the dirt. Then, the laundry had to be put in water again, and the extra water was usually removed to make drying easier. Even if people had water in their homes, the laundry would have to be rubbed against a special board or hit with pieces of wood to make it clean. It was hard work that took a long time.

The first washing machines were operated by hand, and they still required a lot of hard work. Discovering how to use electricity to power these machines was a challenge because the combination of water and electricity is very dangerous. However, during the first half of the 20th century, inventors created electric machines that were able to automatically do most of the steps involved in washing. Before long, these machines became common in homes wealthier parts of the world.

Automatic washing machines gave people more time and energy for other activities than any other new technology did. They used some of this extra time and energy to study and teach their children. This, in turn, led to improvements in the quality of everyone's lives in the places where washing machines became common. Even today, many people in the world still wash their clothes by hand. This means that, over the next few decades, washing machines will probably continue to make a big difference to the lives of billions of humans.

[問題]

(1) What do some historians say about the invention of washing machines?
1. It happened due to an important change in society.
2. It led to the development of TVs, computers, and smartphones.
3. It has had a major impact on the natural environment.
4. It has had a greater effect on society than other modern inventions.

(2) Cleaning clothes and sheets without washing machines was hard work because
1. the process of doing laundry involved several different stages.
2. the soap used to wash laundry had to be prepared by hand.
3. people had to travel long distances in order to dry their laundry.
4. people who did it had to wash many items to earn enough money.

(3) What was one challenge faced by people trying to invent electric washing machines?
1. Many people thought that they would not be as effective as washing laundry by hand.
2. The use of electricity was limited to a few homes in wealthier parts of the world.
3. They could not discover how to make a machine to do all the steps involved in washing.
4. Machines that involve both electricity and water can be very unsafe to work with.

(4) Washing machines have allowed people to
1. spend more time teaching themselves and their children.
2. use their energy for volunteer activities in their communities.
3. invent other machines to carry out tasks in the home.
4. live in parts of the world where there are many rivers and lakes.

The Evolution of Laughter

Laughter is not only a way to express our feeling that something is funny, but it is also something that is good for our health. In the short term, it can help to relax muscles and improve blood flow, and in the long term, it can make our bodies better at fighting diseases. Researchers have been investigating how laughter evolved in humans by looking for similar behavior in other animals. A study carried out at the University of California, Los Angeles, has revealed evidence of laughter-like behavior in over 60 species.

It has long been known that chimpanzees laugh, although the sound is a little different from human laughter. When most humans laugh, they only make a noise when they breathe out, but when chimpanzees laugh, they make a noise both when they breathe out and when they breathe in. Chimpanzees are closely related to humans, so it is not really surprising that they, gorillas, and orangutans laugh. However, as these animals do not have the complicated languages needed to tell jokes, the researchers were interested to find out what makes them laugh.

The researchers found that chimpanzees made these laughing noises when they were playing roughly with each other. They believe that laughter is a chimpanzee's way of letting others know that it is not really trying to harm them. Playing allows chimpanzees and other animals to develop fighting and hunting skills as well as to build stronger relationships with the other members of their groups.

By listening for the noises made by other animals during play behavior, the researchers were able to identify "laughter" in a wide range of animals. Dogs, for example, breathe loudly when they play, and dolphins make special clicking noises. In the case of rats, the laughter-like sounds they make when they are touched gently are too high for humans to hear. However, the sounds can be detected with special equipment. The researchers have concluded that laughter began to evolve as a signal to others that they can relax and have fun. Of course, humans laugh for a variety of reasons, so researchers still have much to learn about how this behavior evolved.

[問題]

(1) How are researchers trying to find out about the development of laughter in humans?
1. By searching for behavior that seems like laughter in other species.
2. By analyzing the kinds of things that people think are funny.
3. By studying the reactions of human babies from the time they are born.
4. By investigating the muscles that are used when a person laughs.

(2) How is chimpanzees' laughter different from most humans' laughter?
1. Chimpanzees make the same noises as humans do when they are surprised.
2. Chimpanzees produce sounds by breathing through their noses.
3. Chimpanzees do not only make sounds when they breathe out.
4. Chimpanzees do not breathe as slowly as humans do when they laugh.

(3) Researchers think that chimpanzees use laughter to
1. indicate that their behavior is not serious.
2. welcome new members to their groups.
3. warm their muscles up before they go hunting.
4. avoid fighting by scaring other chimpanzees away.

(4) Special equipment needs to be used in order to
1. measure the signals in humans' brains when they laugh.
2. recognize the different noises made by dolphins.
3. observe the laughter-like noises of a kind of animal.
4. identify the exact reason that a human is laughing.

資料：熟達度テスト[3]

Desert Delight

The Tohono O'odham people are Native Americans who come from the Sonoran Desert. In fact, the name of this tribe means "desert people" in their own language. The Sonoran Desert lies around the border between the United States and Mexico. Traditionally, the Tohono O'odham people lived in villages and grew crops such as beans, corn, and melons. They also ate some of the wild plants and animals that are found in the desert.

Although the Sonoran Desert is hot and dry, it has over 2,000 different species of plants. Hundreds of these plants are safe for people to eat. There are two reasons why the Sonoran Desert has so many species of plants. One is that it contains a variety of types of soil, and these support the growth of many kinds of plants. The other is that, although the desert is mostly dry, it rains a couple of times each year once in the winter and once in the summer. This rain is enough for some kinds of plants to survive.

One desert plant, the saguaro cactus, is especially important to the people of the Tohono O'odham tribe. Saguaro cactuses can live for over 200 years and grow more than 15 meters tall. Once a year, around June, they produce red fruit. This fruit the saguaro fruit has long been a favorite food of the Tohono O'odham people. When the fruit is ready to eat, families work together to knock it down from the cactuses and collect it. The fruit is sweet and delicious when it is fresh, and it can also be turned into sauce or wine so that it can be stored for long periods.

The people of the Tohono O'odham tribe were very independent, and for a long time, they fought to keep their traditional way of life. However, in the early 20th century, the U.S. government forced them to change their lifestyle. It sent Tohono O'odham children to schools to make them learn English and forget their own culture. Many stopped following their traditional way of life. Recently, though, some Tohono O'odham people have begun bringing back their tribe's endangered traditions, including collecting and eating saguaro fruit.

[問題]

(1) What is true about the Tohono O'odham people of North America?
1. They used to protect the border between Mexico and the United States.
2. They lived in small communities and kept farms in a dry area.
3. They ate wild plants and animals instead of growing their own food.
4. They were forced to leave their homes and live in the Sonoran Desert.

(2) What is one reason that over 2,000 different types of plants can survive in the Sonoran Desert?
1. The sunshine in the area means that some plants can actually grow better there.
2. The Sonoran Desert gets enough rain twice a year to allow the plants to grow.
3. There are few human beings or wild animals living in the region that eat them.
4. There is one kind of soil in the desert that almost any plant can grow in.

(3) The saguaro cactus
1. produces fruit that the local people have enjoyed for a long time.
2. was discovered by the Tohono O'odham people about 200 years ago.
3. has roots that grow 15 meters below the ground to reach water.
4. is best to eat with a special sauce made from traditional wine.

(4) Why did many Tohono O'odham people stop following their traditions?
1. The U.S. government wanted them to behave more like other U.S. citizens.
2. The U.S. government offered them opportunities to travel overseas to study.
3. They wanted their children to study English so that they could enter good schools.
4. They lost their independence after a war that took place in the early 20th century.

Competing against Braille

Although Braille is the standard writing system for blind people today, this alphabet of raised dots representing letters was not always the only system. Another system, Boston Line Type, was created in the 1830s by Samuel Gridley Howe, a sighted instructor at a US school for blind people. Howe's system utilized the letters in the standard English alphabet used by sighted people, but they were raised so they could be felt by the fingers. Blind students, however, found it more challenging to distinguish one letter from another than they did with Braille. Nevertheless, Howe believed that the fact that reading materials could be shared by both blind and sighted readers outweighed this disadvantage. His system, he argued, would allow blind people to better integrate into society; he thought Braille encouraged isolation because it was unfamiliar to most sighted people.

It gradually became clear that a system using dots was not only easier for most blind people to read but also more practical, as the dots made writing relatively simple. Writing with Boston Line Type required a special printing press, but Braille required only simple, portable tools, and it could also be typed on a typewriter. Still, despite students' overwhelming preference for Braille, Boston Line Type remained in official use in schools for the blind because it allowed sighted instructors to teach without having to learn new sets of symbols. Even when Boston Line Type lost popularity, other systems continued to be introduced, leading to what became known as the "War of the Dots," a situation in which various writing systems competed to become the standard.

One of these, called New York Point, was similar to Braille in that it consisted of raised dots. Its main advantage was that typing it required only one hand. Braille, though, could more efficiently and clearly display capital letters and certain forms of punctuation. There were other candidates as well, and debates about which was superior soon became bitter. Blind people, meanwhile, were severely inconvenienced; books they could read were already in short supply, and the competing systems further limited their options, as learning a new system required great time and effort. At one national convention, a speaker reportedly summed up their frustrations by jokingly suggesting a violent response to the next person who invents a new system of printing for the blind.

The War of the Dots continued into the 1900s, with various groups battling for funding and recognition. In the end, the blind activist Helen Keller was extremely influential in ending the debate. She stated that New York Point's weaknesses in regard to capitalization and punctuation were extremely serious and that reading it was hard on her fingers. Braille won out, and other systems gradually disappeared. Although the War of the Dots interfered with blind people's education for a time, it had a silver lining: the intense battle stimulated the development of various technologies, such as new typewriters, that greatly enhanced blind people's literacy rates and ability to participate in modern society.

資料：熟達度テスト[4]

[問題]

(1) What did Samuel Gridley Howe believe about Boston Line Type?

1.The time it saved blind people in reading made up for the fact that it took much longer to write than Braille.
2. The fact that it combined raised dots with other features made it easier for blind people to use it when communicating with one another.
3. Although it was difficult for students to learn, the fact that it could be read more quickly than Braille was a major advantage.
4. It was worth adopting because of the role it could play in helping blind people to better fit in with people who are able to see.

(2) In the second paragraph, what does the author of the passage suggest about Boston Line Type?

1. Its continued use was not in the best interests of blind people, whose opinions about which system should be used were seemingly not taken into account.
2. Teachers at schools for the blind convinced students not to use it because they thought systems with fewer dots would be easier for students to read.
3. Despite it causing the "War of the Dots," its popularity among students was a key factor in the development of other tools for blind people.
4. It was only successfully used in writing by students in schools for the blind after the introduction of the typewriter.

(3) The suggestion by the speaker at the national convention implies that blind people

1. felt that neither Braille nor the New York Point system could possibly meet the needs of blind readers.
2. were unhappy that the debates over which system to use were indirectly preventing them from accessing reading materials.
3. did not like that they were being forced to use a writing system that had not been developed by a blind person.
4. were starting to think that other types of education had become much more important than learning to read books.

(4) What conclusion does the author of the passage make about the War of the Dots?

1. It was so serious that it is still having a negative influence on the research and development of technology for the blind today.
2. It would have caused fewer bad feelings if Helen Keller had not decided that she should become involved in it.
3. It had some positive effects in the long term because the competition led to improvements in the lives of blind people.
4. It could have been avoided if people in those days had been more accepting of technologies like the typewriter.

The British in India

Established in 1600, the British-owned East India Company was one of the world's largest corporations for more than two centuries. By trading overseas with various countries, such as India and China, it was able to import luxury items from these countries into Britain. The British government received a portion of the company's vast profits, so it was more than willing to provide political support. Due to its size, power, and resources, which included a private army of hundreds of thousands of Indian soldiers, the company pressured India into accepting trade contracts that, in general, were only of benefit to the company. After winning a battle against a local ruler in the 1750s, the company seized control of one of the wealthiest provinces in India. As a result, the East India Company was no longer solely acting as a business but also as a political institution, and it began forcing Indian citizens to pay it taxes.

The East India Company gained a reputation among the countries it did business with for being untrustworthy. It also started to lose popularity within the British Parliament because the company's dishonest trading habits damaged foreign relations with China. Then, in the 1850s, angered by the way they were being treated, a group of soldiers in the East India Company's army rebelled. They marched to Delhi to restore the Indian emperor to power, and their actions caused rebellion against the British to spread to other parts of India. The rebellion was eventually brought under control after about two years, but it triggered the end of the East India Company. The British government, which blamed the East India Company for allowing the rebellion to happen, took control of India, and an era of direct British rule began. The British closed down the East India Company, removed the Indian emperor from power, and proceeded to rule India for almost a century.

While some claim that India benefited from British rule, typically using the construction of railways as an example, many historians argue that the country was negatively affected. In an effort to reinforce notions that British culture was superior, Indians were educated to have the same opinions, morals, and social preferences as the British. The British also implemented a policy known as "divide and rule," which turned Indians from different religious backgrounds against each other. The British government used this strategy to maintain its control over India, as members of these religions had joined forces during the earlier rebellion. However, nationalist feelings among Indians increased from the early 1900s, and India eventually gained its independence in the late 1940s.

Although the East India Company stopped operating more than a century ago, it has had a lasting influence. Some experts say it pioneered the concept of multinational corporations and ultimately led to the economic system of capitalism that is widespread today. Moreover, the connection between the British government and the East India Company set a precedent for using political power to help achieve business objectives.

資料：熟達度テスト[5]

[問題]

(1) What was one result of India doing business with the East India Company?

1. India could afford to increase the size of its military because it was able to make trade deals with other countries.
2. India had little choice but to agree to business agreements that were unfavorable to it.
3. The Indian government needed to raise taxes in order to pay for losses from failed trade contracts.
4. The Indian government's relationship with China became worse, which almost resulted in a break in trade between the two countries.

(2) What led to the British government taking control of India?

1. The British government held the East India Company responsible for an uprising that occurred.
2. The Indian people voted for British rule after losing confidence in the Indian emperor's ability to rule the country effectively.
3. The Indian people asked for the help of the British in preventing a war between India and China.
4. The Indian emperor decided to join forces with the British as a political strategy to maintain control of India.

(3) One effect that British rule had on India was that

1. Indians were able to take part in the process of building a government that reflected their economic and social needs.
2. schools made an effort to educate their students to have an awareness of both Indian and British cultures.
3. divisions were created between different groups of Indians to prevent them from challenging British rule.
4. many of the railroads and other transportation systems built by the Indian government were destroyed.

(4) What does the author of the passage say about the East India Company?

1. The company prevented the British government from achieving its aim of expanding its rule to other countries in Asia.
2. While the company may have been successful during its time, its business model would not be effective in today's economy.
3. Although the company no longer exists, it has had a large impact on the present-day global economic landscape.
4. If the company had never been established, another one would likely have ended up having similar political and economic influence.

Animal Talk

How do you tell a friend if you are excited? Most animals communicate using their voices or bodies.

Howler monkeys are the loudest animal in the world except for the blue whale. Howler monkeys live in the tropical rainforests of South America in groups of fifteen to twenty monkeys. Adult male Howler monkeys are black or brown and female or younger monkeys are light brown. Every morning and evening, Howler monkeys have a group howl, a noise that can be heard three miles away. This loud howl keeps other animals away from the troops' food source. Howler monkeys have big necks and lower jaws because they have extra-large vocal cords.

The Black Palm Cockatoo, of Australia, uses hollow or dead wood to make noise. These birds are the largest cockatoos and can be 28 inches long.
They have red cheeks, a black and red tongue, and black curving feathers on their back. Because they do not have loud voices, the black palm cockatoo and its mate use small *twigs in their claws to tap on a hollow tree. This warns other animals to stay away from their territory.

Some animals make sound by rubbing body parts together. Spiny lobsters, or crayfish, make a rasping sound like a finger being rubbed along the teeth of a comb. This noise keeps the group of lobsters together and signals that the group is safe. If a shark or eel shows up for a meal, the sound speeds up and becomes higher. Spiny lobsters are red-colored and look similar to lobsters. They have thick, spiky antennae and do not have claws.

*twig: 小枝

資料：読解テキストB

The Rodeo

The first rodeos began in the 1800s as a way for working cowboys to show off their riding and roping skills. Today's rodeos have many of the same events as the early rodeos.

In the Saddle Bronc event, a cowboy rides a bucking horse. The rider holds a thick rein with one hand and tries to stay on the horse. The rider receives points for the position of his toes and how much control he has over the horse. On the range, working cowboys captured wild horses and *tamed them for their work.

The Saddle Bronc rider must stay on the horse at least 8 seconds. The rider is disqualified if he touches any part of the horse or his own body with his free hand. Many people consider Saddle Bronc Riding the most difficult sport in rodeo.

In tie-down roping, a rider on horseback chases a calf and lassos it with a 25-foot rope. Next the cowboy dismounts, runs to the calf and ties three of the calf's legs together. When the cowboy is finished tying the calf, he throws his hand in the air. The cowboy is scored on how quickly he lassos and ties the calf. If the calf kicks free within six seconds, the rider receives a score of ‘no time.’

In the 1800s, working cowboys used lassos to catch cattle and to brand them. Barrel racing is one of the few rodeo events open to women.

Barrel racing demonstrates a rider's speed and control when riding quickly. Working cowboys rode on horses to move a large herd of cattle to a new place. The rider rides in a cloverleaf pattern around three barrels and receives points for how quickly she completes the pattern. If a barrel is knocked over, five seconds is added to the total time.

*tame: [動物などを] 飼いならす、手なずける

A 研究部門 ● 報告Ⅱ ● 英語能力テストに関する研究

グラフィックオーガナイザーを用いた心的表象構築プロセスの視覚化と評価

研究者:工藤　大奈　東京都／安田学園中学校高等学校　教諭
（申請時:茨城県／筑波大学大学院　在籍）
《研究助言者:和泉 伸一》

概要　昨年度施行された学習指導要領では，深い学びまでの理解が求められている。深い学びに繋がるために，視覚情報をリーディングに使用することが有効である。今回の調査では，リーディングに，本文の内容を図式化したグラフィックオーガナイザー（Graphic Organizer，以下GO）を取り入れた。現状の英検問題において，GOを作成する問題を設けることが可能なのか，過去問題を精査し調査した（調査1）。加えて，リーディングの最中にGOを作成させるタスクを課し，これにより読み手が作成する読解のイメージを視覚化し，評価までできるのかを調査した（調査2）。結果として，調査1ではメール文と論説文がGO作成に適していると判断された。これは，時系列が明確であったり，複雑な情報の整理が可能であるためである。また調査2では，読み手が作成したGOを評価として活用することによる有益な結果は得られなかった。GOの活用については内容理解のサポートとなることは判明したが，評価の作成については今後の大きな課題である。

1 はじめに

2022年度より高等学校では新たな学習指導要領が施行され，「主体的・対話的で深い学び」が求められている（文部科学省，2018）。新学習指導要領では，あるトピックを自分のこととして捉えて考えることが主体的であると述べられ，さらに「外国語でのコミュニケーションを図る資質や能力を身につける」ためには，主体的で自律的に取り組むことが必要であるとしている。そのため，英文読解を行う際にも，主体的で深い学びは必要になる。主体的に学ぶ姿勢は，生徒たちが大学受験のために学ぶこと以外にも，生涯学習の際の重要なきっかけになると考えられる。

英文読解を行う上で，理解が難しいとなったときに有効なのが，視覚情報である。視覚情報は，文字情報のみで記述されるよりも読み手の注意を引き，内容理解を促進する働きがある（卯城編，2009）。理解に関しては，高等学校外国語学習指導要領の「読むこと」において，「日常的な話題について，使用される語句や文，情報量などにおいて，多くの支援を活用すれば，必要な情報を読み取り，書き手の意図を把握することができるようにする」ことが目標として掲げられている（文部科学省，2018）。さらに，「読むこと」以外でも，「話すこと」「聞くこと」においても，視覚情報は理解を深め，わかりやすくすることに効果的であることが述べられている。

とりわけ英文読解を図式化したグラフィックオーガナイザー（Graphic Organizer，以下GO）を使用することは，テキスト情報を視覚化することが可能であるため，読み手は包括的な読解を行うことができる

(Suzuki, 2006)。英検のように, 限られた時間内でリーディングを行いながら内容理解問題に取り組むためには, 包括的な読解が必要となる。したがって, 英文内容を視覚化したGOは, 受験者にとって設問の処理をする際のサポートツールになると考えられる。

本調査では, 現在の英検の問題にGOを作成する問題を設けることが可能かを調査した。加えて, 設問に取り組みながらGOを完成させることにより, 読み手が作り上げる読解イメージを視覚化できるのかを調査した。前者を調査1とし, 後者を調査2として, 本稿で報告したい。

2 先行研究

2.1 読解における視覚情報の役割

まず, リーディングとは読み手がテキスト情報を受け取り, それを自らの背景知識と統合させて心的表象を構築するプロセスである。心的表象とは, 読み手がテキストを読んだ後に心に残るイメージや理解の構築を指す。読解は読み手とテキストの情報との相互作用によって成り立つものであり, 読み手がテキストと対話をすることで可能になる(Imsa-ard, 2022; Rahmani & Sadeghi, 2011)。Kintsch and van Dijk (1978)は, 読み手が理解するということは, 一貫した心的表象を作ることだと述べており, これはリーディングの過程において重要な要素であると指摘している。つまり, テキストを断片的に理解するのではなく, 読み手の力で情報を統合し, 一貫性のある心的表象を作り上げることで, より深い理解が可能になるという。文章を理解するためには, 読み手が自分の持っている背景知識とテキストに書かれた情報を統合する必要がある(Colliot & Jamet, 2018)。

読み手にとって馴染みのある文章や, 読み手に関連する, もしくはよく知っている分野のテキストの読解は, 深く理解することができる。これは, 読み手の持っている背景知識が影響しているからである(Grabe, 2004)。反対に, 読み手が背景知識として持っていない分野の読解の場合は、理解が困難になる。また, 第二言語・外国語でテキストを読む場合は、言語処理にも大きな負担がかかるため, 一貫した理解が難しくなることがある(石井, 2006)。その際に読み手の補助となるのが, 写真や絵, イラスト, 図といった視覚情報である。その役割として, (1)読み手の興味を引くこと, (2)読む意欲を高めること, そして(3)読み手の理解を促進させることがある(Levie & Lentz, 1982)。視覚情報が読解に有用な理由としては, Paivio (1990)の二重符号化理論が挙げられる。これによると, 文字情報に加えて視覚情報が一緒にインプットされることで情報の記憶される量が増加し, 学習が促進され, 読解に効果的であるとされる。

2.2 グラフィックオーガナイザー

グラフィックオーガナイザーは, テキスト情報を理解するために視覚的に図式化して提示されたものであり, 一般的な提示方法としては, 階層的, 原因と結果, 比較, 循環の4つの描写方法がある(Ellis & Howard, 2005; Praveen & Rajan, 2013)。「階層的」とは, あるトピックについてトップダウン形式で詳細を述べるものである。「原因と結果」とは, 特定の理由によりある結果が生じたことを矢印を使って図式化するものである。「比較」とは1つのトピックに対して賛否を表したり, 対比関係を図式化する。「循環」とは, 物語等の話の流れを矢印で円状に図式化するものである。

GOには次のような役割があり, 読解との関わりについても以下のように述べられている。

(a)読み手にテキストの談話構造(どのようにテキストが構成されているか, トピックセンテンスなどの把握)を意識させること(Jiang & Grabe, 2007)。

(b)読み手が理解を深めるために段落内の情報を整理すること(Praveen & Rajan, 2013)。

(c)トピックに関する重要な情報がどのように構成されているかを示すことで, 情報の伝達や理解を促進す

ること（Ellis & Howard, 2005）。
(d) テキスト理解のために必要な構成と情報を，生徒がもっている個人の経験と関連づけること（Ellis & Howard, 2005）。

2.3 GOとリーディング

GOを使用したリーディングでは，提示するタイミングに関して2つの種類がある。1つは，読み手がテキストを読む前にGOを提示するもの，もう1つはテキストとGOを同時に提示するものである。さらに，GOを使用するリーディングの方法には，3つのパターンがある。1つは読み手がテキストを読みながら部分的に空所となっているGOを完成させていくもの，別にはあらかじめ作成されたGOを見ながらリーディングを行うもの，最後は読み手がテキストを見ながら白紙の状態からGOを作成するものである（Stull & Mayer, 2007）。

2.4 GOを「見る」ことによる学習と「作る」ことによる学習

読み手が予め完成されたGOを見てリーディングを行うことは，GOを「見ること」だけに留まるが，その際にテキストがどのように構成されているのかを理解しようとするため，読み手によるテキストの処理が必要となる。つまり，「見ること」は「見る学習」であると主張されている（Stull & Mayer, 2007）。Stull & Mayer（2007）では，科学に関する説明文を用いて著者が作成した「著者作成型GO群」と，読み手がテキストを読みながらゼロからGOを作成する「学習者作成群」，そしてGOを使用しない統制群と3つのグループ間における理解度の比較が行われた。理解度の測定として短答式問題が課されたが，結果として著者作成群のほうが，学習者作成群よりも有意に高いスコアを示した。そのため，「見る学習」のほうが「作る学習」よりも学びに繋がると解釈された（Stull & Mayer, 2007）。

一方，タイの公立学校10年生で，熟達度の低い64名を対象に行われた実験では，リーディングを行いながらGOを完成させていく群と従来の読解を行う2つのグループを比較して，双方がどの程度読解力を向上させることができるのかを，(a) 多肢選択形式問題，(b) 正誤問題，(c) 記述問題を用いて調査している（Imsa-ard, 2022）。GO活用グループは，複数回のGOの活用法について指導を受けた。結果として，GOの指導を受け，完成させながら読解を行ったグループのほうが，従来の読解を行うグループよりもテストの点数が高く，読解力が高まったことが報告されている。このことから，GOを自ら作成する場合でも，学びが可能であることがわかる。

読み手自身でGOを作成するためには，テキストがどのように一貫した構造となっているかを考える必要があり，学習者自らがテキスト情報を整理しなければならないため，テキストの能動的な学習を促すと考えられる。そのため，テキスト理解に効果的であるだけでなく（Jiang & Grabe, 2007; Stull & Mayer, 2007），生徒の思考力を高めることに寄与できると考えられる（Praveen & Rajan, 2013）。図表の活用に関しては，石井（2006）が，外国語として日本語を学習する学生を対象として，歴史に関する長文の説明文を母語によって筆記再生させるタスクを課した。この研究では，与えられた図表を見ながらテキストを読むグループと，空欄に入るアイデアをテキストから選び出し記入させるという図表完成グループを比較した。その結果，図表の提示は筆記再生において重要なアイデアを選択することをサポートしたが，テキスト全体の再生の促進には効果が見られなかった。また，図表を完成させる場合，重要なアイデアとそうでないアイデアの識別が不十分であり，かつ記入させても単なる書き写しとなったことから，筆記再生に効果的でなかった。この研究結果から，主体的な活動と視覚情報の使用が，テキスト理解にいつも効果的であるとは言えないことがわかる。

Rahmani and Sadeghi（2011）は，英語翻訳を専攻する中級EFL学習者である大学生108名を対象に，ノートテイキングを行うトレーニングを実施し，それによる多肢選択形式と筆記再生テストでの効果を

検証した。この研究では，ノートテイクをGOの一部であると解釈している。事前のGOのトレーニングでは，ノートテイキングの重要性について紹介し，GOとなるノートの取り方のレクチャーが行われた。実験では，文章を読んでノートテイクを行い，その後30分間の30問選択式問題に取り組ませ，文章の覚えていることをすべて書き出すという筆記再生テストが行われた。なお，選択式問題では，参加者はノートテイキングした内容を見直しながら解答することが許された。結果として，ノートテイキングを使用した実験群は，従来通りの読解を行った統制群と比べて，選択式問題及び筆記再生テストの両方で優位な結果が得られた。よって，ノートテイキングを使用するリーディングが，中級EFL学習者の読解と記憶の双方を助けることがわかった（Rahmani & Sadeghi, 2011）。

また，Suzuki（2006）の日本人高校生を対象とした実験では，GO作成群とテキストの要約群における読解方略の比較を思考発話の分類分析によって行った。結果として，白紙からGOを作成する群では，GOによってテキストの整理や視覚化ができたため，読み手への負荷が軽減し，テキスト全体の構成に関わる一般的読解方略（general comprehension strategies）を多く使用することが可能になったと報告されている。

以上の先行研究をまとめると，GOを作成しながら英文読解をすることは，読み手に良い影響を与える場合と，そうでないものに分かれている。対象となった協力者が日本人英語学習者に限らず，属性が様々であり，協力者がGOの使用にどこまで慣れているかといったことも，結果の違いの理由にあると思われる。また、学習者の英語習熟度とテキストのマッチングの問題や、テキストジャンルとGOの相性度など，他にも考慮に入れなければならない要因が考えられる。

3 調査1

3.1 調査の目的

以上の先行研究に基づき，本研究の調査1では，現在英検で出題されているリーディング問題において，GOを作成しやすいものであるのかどうかの調査を行った。調査1における研究課題（Research Question: RQ）は以下の通りである。

RQ1 英検準2級・2級・準1級のリーディング問題で使用されている長文では，どのようなテキストジャンルと設問のタイプがGOの作成に適しているか。

3.2 使用したマテリアル及び方法

調査1では，まず各級で使用されているテキストジャンルと設問タイプを精査し，どのような形式で文章が構成されているのかを検討した。検討したテキストは，過去3年以内に実施された実用英語技能検定の問題から，準2級・2級・準1級のリーディングセクションを選んだ。この選択の理由は，対象を大学生・大学院生に絞ったためである。準2級が高校在学中向け，2級が高校卒業程度向け，準1級が大学中級程度向けのレベルとされているため，中高6年間と大学での必修科目で英語を学習している協力者にとって，適切なレベルであると判断した。

調査の方法としては，各級のリーディングセクションにおける語句の空所補充問題と，内容一致選択問題に焦点を当て，各級に5題含まれているリーディング問題を5回分，すなわち各級計25題のメール文、会話文、説明文読解問題を筆者が1題ずつ解き，GOの提示や作成が可能であるかを判断した。「可能である」と判断したのは，時間軸のイメージや，因果関係を考える必要があり，図式化してテキスト情報の整理をすることが可能になる場合である。この判断は筆者個人によって実施された。

3.3 結果

まず，各級のリーディングセクションの構成についてまとめると，準2級においては，短い文章や簡単な対話形式の問題が多く出題されており，理解力や基本的な表現力が求められる傾向が見られた。2級では，より長文になるに加えて，抽象的な内容の文章が登場し，読解力の向上が必要とされる。準1級では，専門的な知識を要する文章が出題されることがあり，より高度なリーディングスキルが求められる。

今回対象としていた設問の形式として，語句の空所補充問題と内容一致選択問題のどちらも読み取った情報を正確に把握し，適切な判断を下す能力が問われているが，語句の空所補充問題では，文脈を考慮して適切な単語やフレーズを選ぶことが求められていた。このような問題に対応するためには，文章全体の流れを把握し，論理的なつながりを理解する能力が重要になると思われる。

各級における文章構成の違いについては，準2級に関しては以下の通りである。当該級のリーディングセクションでは、語句の空所補充をするものが2題小問計5問，メール文読解が1題小問3問，内容理解問題が1題小問4問で構成されていた。2級については，語句の空所補充をするものが2題小問計6問，メール文の内容理解問題が1題小問3問，論説文の内容理解問題が2題小問計9問となっていた。次に準1級においては，語句の空所補充をするものが2題小問計6問，論説文の内容理解問題が3題小問計10問という構成になっていた。

各級における設問のタイプとしては，語句の空所補充問題では，文脈から推測が必要になるものも見られるが，GOを作成した上で解答するには適していないと判断された。これは，文脈に合う語句を選択することが，必ずしもテキスト全体の理解と同じであるとは言えず，記憶力テストとして空所補充問題が使われている例が見られたためである（Stull & Mayer, 2010）。

次にテキストジャンルとしては，メール文と論説文が特にGO作成には適していると判断された。メール文は論説文とは異なるが，時系列を示す語句が使われていることから，時間軸のイメージをGOの作成がサポートできると考えた。これは，Suzuki（2006）で使用されていたマテリアルにおいて，時系列を意識したという点で一致している。論説文においては，GOを活用することで，ややもすると複雑となりがちな情報の整理を視覚的に行えると考えられる。複数のGOを作成させる先行研究でも，科学や心理学に関連する論説文が使用されており，英検の論説文でも使用できると判断した（Praveen & Rajan, 2013）。なお，準1級ではメール文は使用されていなかったため，メール文に関しては2級と準2級のみでの判断となる。また，会話文は準2級のみであったため，今回の調査対象から除外した。

以上のことから，次の調査2では，メール文と論説文を活用して生徒にGOを作成させたのち，内容一致選択問題に取り組むというタスクを課した。

4 調査2

4.1 方法

調査2では，調査1で得られた情報を基に，GOを使用した内容理解の評価が可能であるかどうかを明らかにすることを目的として実施した。使用したGOは，調査1でメール文や論説文で見られた時系列や矢印を使用して因果関係や対比関係を説明できるものとした。

調査2では，以下のように2つのパターンで実験を行った。まず，同じ協力者を対象に，(1)GOを使用せずに従来通りの読解問題4題（内容一致選択問題14題）に取り組んでもらい，その後で(2)GOの作成を伴う別の読解問題4題（内容一致選択問題14題）に取り組んでもらった。これにより，GOをリーディングに導入することで内容理解にどのような影響があるのかや，GOの作成によって読み手が読解でイメージしたものを視覚化できるかを検証した。研究課題は以下の通りである。

RQ2　リーディングにおけるGOの作成タスクは，評価の一種として成立するか。

4.2 協力者

本調査での協力者は，国立大学に通う大学生及び大学院生15名であり，女性8名，男性7名であった。専攻は，人文社会・理工学・医学・看護・体育・生命環境と多岐にわたり，中学校及び高等学校で少なくとも6年間の学校英語教育を受けてきた。協力者への事前アンケートの結果から，所有している英語資格として，実用英語技能検定2級から準1級，TOEIC L & Rの結果において545点から870点であることがわかり，ヨーロッパ言語共通参照枠（CEFR）においてA2からB2に当てはまることがわかった。

4.3 マテリアル

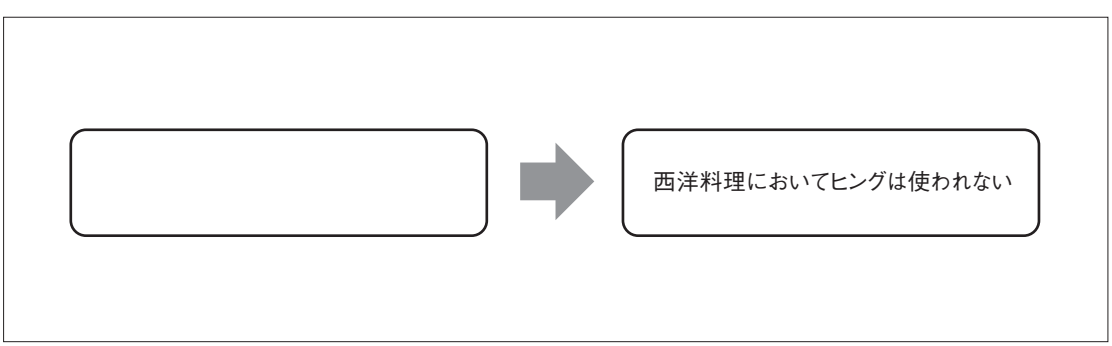
4.3.1 使用したリーディングテキスト

使用したリーディングテキストは，英検準2級・2級・準1級のリーディングセクションである。課題（1）と（2）ともに2級のメール文，2級の説明文，準2級の説明文，準1級の説明文の順番で，計4題14問の小問に取り組んでもらった。このようにマテリアルを設定した理由としては，高等学校のレベルである準2級から大学中級程度のレベルである準1級とレベルの幅広いマテリアルに触れることで，協力者のレベルに偏らずに調査できるためである。また大学生・大学院生を対象としたこともあり，全員が高等学校卒業程度の学力は備わっていると見て，2級のテキストを1題多く用意した。メール文と説明文の両方を使用した理由としては，GOが異なるジャンルのテキストで使用可能かを探るためである。

4.3.2 使用したGO

GOを使用するリーディングの方法には，3つのパターンがあることは前に触れた（読みながら部分的に空所を補充するもの，あらかじめ作成されたGOを見ながらリーディングを行うもの，白紙の状態からテキストを見ながらGOを作成するもの）。今回は，協力者が十分にGOに馴染みがあったわけではなかったため，部分的に完成させていく形式を採用した。

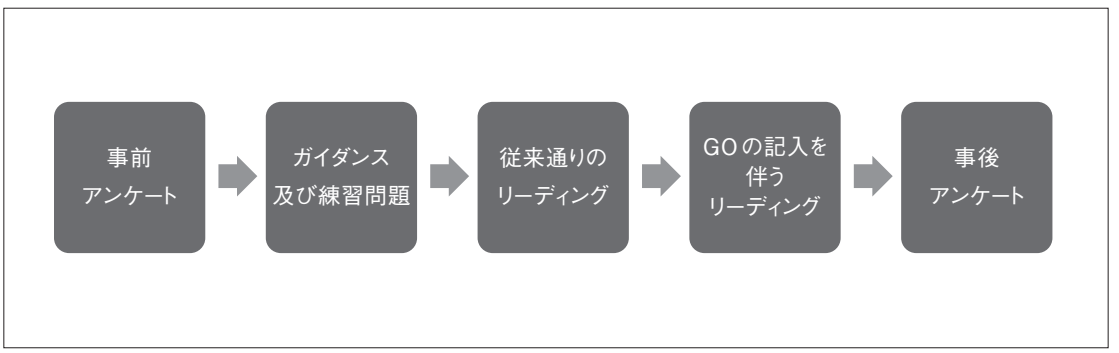
本調査で使用したGOは，原因と結果や対比関係，時系列を矢印の関係で提示したものである。英検の説明文やメール文では，1つの段落につき1つの小問のように構成されているため，内容理解問題の手助けになるように，1つの小問につき1つのGO（1つの矢印の関係）を提示するようにした。例えば，説明文において，特定の国で使われる調味料が西洋料理では使われない理由に関する説明がなされていた場合（例えば“It is not often used in Western food, probably because not many cooks know about it.”），結果部分を予め埋めて置き，原因部分を空白として，協力者にそこを日本語で埋めてもらった（図1）。



■図1：調査2で使用されたGOの例

4.3.3 手順

本調査の手順は図2の通りである。



■図2: 調査2の手順

事前アンケートは協力者が所有している英語資格や留学経験の有無を尋ねるため，及びGOを使用することについて意見や感想を自由に述べてもらうために実施した。ガイダンスでは，調査の目的や，GOの記入である空所補充の方法について説明し，準2級のリーディングセクションにおけるメール文を使用して，練習問題を実施し，実施に際しての不明点を解消してもらった。また実施時間については個人差があるものの，平均して90分程度であった。なお，実施方法としては，オンラインで行い，GOの作成過程も調査するために，実験中は協力者の許可を取って録画・録音をした。

4.4 分析と結果

4.4.1 読み手の心的表象の視覚化について

本調査では，協力者の読解中の思考回路や考え方，答えの導き方についても視覚化させるために，リーディングをしながら頭に思い浮かんだことや考えたことを発話してもらうという思考発話調査も行ったが，当初想定したほど発話がうまく得られず，且つ音声ファイルのデータ保存の技術的問題があったため，本報告の分析からは除外することとした。

協力者が完成させたGOは，以下のようなルーブリック評価を行った。ルーブリックは筆者が作成したのち，観点について英語教育学を専攻する大学院生に確認をしてもらった。疑問点等は話し合ったのち，基準となる評価を決定し使用することとした(表1)。各協力者のGOルーブリック評価は，上記と同じ大学院生に協力してもらった。評価方法としては，GOありのリーディング問題において，4題ある大問1つずつに対してA・B・Cと評価をつけ，4題の評価の結果から，総合評価A・B・Cをつけて各協力者への評価を決定した。全協力者の解答のうち30%を上記の大学院生に評価してもらい，不一致の箇所については協議したのち，残りの協力者の解答は筆者が実施した。本調査では，ルーブリックでの評価を3段階の観点から算出し，あわせて内容理解問題の設問正答率を算出した。GOなしとありの場合の結果を，表2と表3にそれぞれ示す。

■表1: 使用した評価のためのルーブリック表について

観点	A	B	C
GOの空所補充と設問解答の根拠となる要点の把握ができている。	テキスト及び前後の内容にあわせて正しくGOの作成ができ，要点を正しく把握している。(完答)	概ね正しくGOが完成できていて，要点の把握ができている。(空欄2箇所が正しい)	内容にあったGOの完成ができていない。(空欄1箇所が正しい，もしくは全不正解)

■表2: GOなしの場合の設問正答率

協力者	設問正答率%
1	71.4
2	92.9
3	100.0
4	92.9
5	92.3
6	85.7
7	100.0
8	92.9
9	71.4
10	78.6
11	92.9
12	71.4
13	92.9
14	92.9
15	92.9

■表3: GOありの場合の設問正答率とルーブリック評価

協力者	ルーブリック評価	設問正答率 %
1	B	78.6
2	B	71.4
3	B	78.6
4	B	85.7
5	B	85.7
6	B	92.9
7	C	71.4
8	B	78.6
9	B	64.3
10	B	100.0
11	B	92.9
12	B	78.6
13	B	71.4
14	B	85.7
15	C	78.6

上記の表についてそれぞれを述べていくと，表2については，協力者にGO を与えず従来通りの読解を行い，14個の内容理解問題に取り組んでもらった際の正答率である。表3については，GO作成しながら読解を行い，14個の内容理解問題に取り組んでもらった際の正答率と，作成したGOを表1のルーブリックを使用して評価した結果である。ルーブリック評価と内容理解問題の正答率の関係性について相関関係を調べると，ピアソン相関係数は0.13となった。本調査の結果では，ルーブリックによる評価にて尺度がBとCのみ，特に評価がBに偏る結果となった。また，GOの有無による内容理解問題への影響を見るために*t*検定を実施したところ，$p>.05$ となり，両者に有意差は見られなかった。つまり，GOの作成が内容理解問題への正解を促進する結果とはならなかった。

4.4.2 Googleフォームでのアンケート結果

本調査では，リーディングのあとに，次の質問に答えてもらった:「内容理解問題の前に下線部を埋める作業がありました。これについて意見や感想を自由に述べてください」。表4に，協力者の回答と，内容理解問題の正答率，及びGOの評価をまとめて記す。結果として，15名中10名がGOに対して肯定的な意見を述べていた。参加した協力者の多くがGOで空所補充を行うことが内容理解問題の助けになったと回答している。加えて，空所補充を行うことによって、リーディングの内容をより良く記憶できたという意見もあり，内容理解問題のサポートになっていたと考えられる。ただし，中には否定的な意見を述べた協力者もおり，「面倒くさい」や「文脈が分かりにくかった」という感想を持っていた。

■表4: 事後アンケート結果 (原文ママ)

協力者	アンケート回答	内容理解問題正答率	GO ルーブリック評価
1	文章の内容がよくつかめていなくても，ヒントとして使うことができたため，問題を解くのに役だったのではないかと思う。	78.6	B
2	日本語の接続がわかりづらく，何と答えてよいかわからない問題がいくつかあった。	71.4	B
3	下線部を埋める形式よりは単に「〜なのはなぜか?何か?」みたいな設問にして自由に記述できる形式の方がやりやすいと感じた。	78.6	B
4	自分が分かる単語で構成された文章に関しては，2セット目の方が一度日本語に翻訳したものを可視化することができ，問題を解きやすいと感じた。しかし，分からない単語が多い文章の場合には，日本語に翻訳する作業で混乱してしまい，時間がかかったように感じた。また，内容理解問題に影響はなかったが，下線部にどこまでの内容を書けばよいのかが分からないこともあった。	85.7	B
5	場合によっては下線部を埋める作業と内容理解問題を並行して進めましたが，内容の理解を助ける効果がこの作業にはあったと感じました。	85.7	B
6	自分が誤解していたり読み飛ばしていた箇所をもう一度読みなおして英文を正しく認識することができた。	92.9	B
7	めんどうだった。	71.4	C
8	下線部を埋める作業によって，してないときよりも文章をより理解でき，読解した内容を覚えていた。	78.6	B

協力者	アンケート回答	内容理解問題正答率	GO ルーブリック評価
9	日本語で要約する形式だったので，日本語で内容把握ができて，スムーズに行うことができたように感じた。	64.3	B
10	文脈がややこしいことがあった	100.0	B
11	読解のヒントになっていて比較的問題を解きやすかったです。	92.9	B
12	穴埋めだったため，他の部分で日本語が多く，それをヒントにして読解を進めていくことがあった。穴埋めがあると，全体の文章を先に把握できるため，読解が簡単なように思えた。	78.6	B
13	下線部を埋めようと意識して読解をすると，より本文の内容が頭で整理されやすくなると感じた。	71.4	B
14	面倒だった。	85.7	B
15	文章を理解しやすくなった。	78.6	C

5 総合考察

調査1においては，「英検準2級・2級・準1級のリーディング問題で使用されている長文では，どのようなテキストジャンルと設問タイプがGOの作成に適しているか」という研究課題のもと，リーディングセクションについて検証した。結果として，内容一致選択問題におけるメール文と説明文がGOの作成に適しており，理解に繋がるGOの作成や提示が適していると判断された。とりわけメール文ではGOを使用することで情報の視覚化や階層化を助け，説明文では内容を論理的により分かりやすくする上で役立つのではないかと考えられた。例えば，複雑なプロセスを説明する場合には，矢印を伴うフローチャートやタイムラインを明示するGOを用いて，直感的に理解しやすい形式にまとめることができる。学術論文や専門的なテーマを扱う場合であっても，専門用語の整理や重要なポイントの強調をGOによって行うことで，読者の理解を助けることも可能であると思われる。

調査2では，「リーディングにおけるGOの作成タスクは評価として成立するか」という研究課題のもと検証を行った。相関係数やt検定の結果から，ルーブリックによる評価は本調査では有意な結果は見られなかった。そのため，今回の結果からは，ルーブリックの評価と設問の正答率には明確な関係があるとは言うことができなかった。つまり，設問正答率が良いからと言ってGO作成の力があるとは言えず，また，GO作成の力が備わっていないからと言って内容理解問題が解けないということにはならない。

次に，読み手が読解中に作り上げる心的表象について述べる。まず心的表象の作成という点については，矢印の前後を埋めるために，長文から必要な情報の取捨選択を行う必要があった。作成過程を録画したものを確認すると，過半数の協力者は，一度で必要な情報を見抜き，正しい解答を矢印の前後に書き込むことができたが，一部の協力者は一度自分の答えを書き込んだのち修正し，正しい解答に書き直していたのが見られた。つまりこの際に，後者の協力者は，作り上げた心的表象の修正を行っていたと考えられる。この過程では，「自分の答えを書く→誤りに気付く→修正して正しい答えを書く」という3つのステッ

プを見ることができた。テキストで描かれている状況の把握をすぐにできた協力者は，GOを作成する際に矢印の前後の空所に正しく記入し，内容理解問題に解答していた。よってGOの作成が内容理解問題に解答する際のサポートとなり思考が促進され，正しく解答できたのではないかと考えられる。一方，状況の把握ができなかった協力者は，GOの作成もうまくできなかったことが考えられる。

次に，調査2で実施したアンケート結果について考察する。まず，各協力者の意見とそれぞれの内容理解問題の正答数との関係について述べる。協力者1は，GOをヒントとして活用したことから内容理解でも役立ったと述べており，ルーブリック評価でも概ね良い結果を出せていた。他の協力者も同様な回答と評価及び正答率が見られるが，協力者6については，一概にGOの効果があったとは言えない。日本語で書かれたGOと英文読解での理解に齟齬が出た可能性もあると考えられる。また，「文章を先に把握できたため，読解が簡単なように思えた」と述べているが，協力者12が描いた心的表象と，ヒントとして提示したGOに差異が出ていた可能性がある。さらに，協力者8は，「下線部を埋める作業によって、してないときよりも文章をより理解でき、読解した内容を覚えていた」と述べていながらも，内容理解問題の正答率やルーブリック評価の結果は芳しくなかった。文章理解と問題解決が異なる課題となってしまっていたことが考えられる。

これらの結果を踏まえると，GOを活用することは，一部の協力者には内容理解に対してサポートとなったが，その効果は一様ではなかった。アンケート回答の肯定的･否定的な意見についても個人差が見られ，GOの効果は全般的に個人差があることがわかった。以上に鑑み，リーディング問題の改善や，学習支援のあり方に関して，更なる研究が必要となると言える。加えて，今後は，ルーブリック評価の観点をより細かくして，判定に偏りが出ないような評価を作成することも課題として考えられる。アンケート結果に示された意見を反映しつつ，多様な学習スタイルに対応できる教材やアプローチを構築し，学習効果の向上を模索することが今後の課題である。

6 まとめ

今回の調査では，現状の英検リーディング問題におけるGO使用の可能性と，リーディングをしながら読み手がGOを作成した場合の読み手へのルーブリック評価の可能性について調査した。結果として，作成したGOを評価の一環として活用することにおいては，あまり有益な結果は見られなかった。また，GOの活用が全ての学習者にとって有効であるとも言い切れない結果となった。よって，今後の調査では，以下の点について考慮する必要がある。まず，個人差の影響である。学習者間には認知能力や学習方法に差があると想定されることから，GOの効果がそれらの個人差によって異なってくる可能性がある。今後の研究では，個人の特性を詳細に分析することで，GOの効果に影響を与える要因を把握することが必要となると考えられる。加えて，GOのデザインや使いやすさの向上，またGOを活用したリーディングの練習も必要になるだろう。さらに，小説等の物語文におけるGOの可能性についても，今後調査する必要があるだろう。物語文は，英検で出題が見られなかったため，今回の調査では研究対象とはしなかったが，その実質的な重要性から，今後GOとの関係について掘り下げていくことが求められるだろう。

謝辞

今回このような研究と執筆の機会をいただけましたことに感謝いたします。選考委員の先生方，公益財団法人 日本英語検定協会 事務局の方々，とりわけ最後までとても丁寧にご指導していただきました和泉伸一先生に，心より御礼申し上げます。また，英検研究助成に応募をするにあたって指導をしてくださいました筑波大学の卯城祐司先生，採点や評価をするにあたってサポートをしてくださいました，筑波大学大学院英語教育学サブプログラム博士後期課程1年の水書亮さんに感謝いたします。

引用文献

Colliot, T., & Jamet, É. (2018). Does self-generating a graphic organizer while reading improve students' learning? *Computers & Education, 126*, 13–22. https://doi.org/10.1016/j.compedu.2018.06.028

Ellis, E., & Howard, P. (2005). Graphic Organizers: Power Tools for Teaching Students with Learning Disabilities. *Learning Disabilities, 6*. Retrieved from https://lincs.ed.gov/professional-development/resource-collections/profile-149

Grabe, W. (2004). Research on teaching reading. Annual Review of *Applied Linguistics, 24,* 44-69. Retrieved from https://www.cambridge.org/core/journals/annual-review-of-applied-linguistics/article/abs/3-research-on-teaching-reading/6B9BEBF2442C3850641F70F0F12B3C19

石井怜子. (2006). 図表の呈示及び完成が第二言語学習者の説明文読解に及ぼす影響 中級後半レベルの成人日本語学習者の場合. 教育心理学研究, 54(4), 498-508.

Imsa-ard, P. (2022). Reading Better?: Enhancing Thai EFL Secondary School Students' Reading Comprehension Abilities with the Use of Graphic Organizers. *English Language Teaching, 15*(5). https://doi.org/10.5539/elt.v15n5p1

Jiang, X., & Grabe, W. (2007). Graphic organizers in reading instruction: *Research findings and issues.* Retrieved from http://hdl.handle.net/10125/66616

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review, 85*(5), 363.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *ECTJ, 30*(4), 195-232. https://doi.org/10.1007/BF02765184

文部科学省. (2018).『学習指導要領(平成30年告示)解説　外国語編』.東山書房.

Paivio, A. (1990). *Mental representations: A dual coding approach.* Oxford University Press.

Praveen, S. D., & Rajan, P. (2013). Using Graphic Organizers to Improve Reading Comprehension Skills for the Middle School ESL Students. *English Language Teaching, 6*(2), p155. https://doi.org/10.5539/elt.v6n2p155

Rahmani, M., & Sadeghi, K. (2011). Effects of note-taking training on reading comprehension and recall. *Reading, 11*(2), 116128. Retrieved from https://readingmatrix.com/articles/april_2011/rahmani_sadeghi.pdf

Stull, A. T., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology, 99*(4), 808-820. https://doi.org/10.1037/0022-0663.99.4.808

Suzuki, A. (2006). Differences in reading strategies employed by students constructing graphic organizers and students producing summaries in EFL reading. *Jalt Journal, 28*(2), 177. Retrieved from https://jalt-publications.org/files/pdf/jalt_journal/2006b_jj.pdf#page=69

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.研究社.

A 研究部門 ● 報告Ⅲ ● 英語能力テストに関する研究

英字幕付き音声動画の視聴による第二言語の偶発的語彙学習

― 使用する音声動画マテリアルに関わる変数が持つ影響 ―

研究者:黒川　皐月　宮城県／東北大学大学院　学術研究員
（申請時:東京都／東京外国語大学大学院　在籍）
《研究助言者:西垣 知佳子》

概要

本研究は，英語を外国語として学習する日本人大学生148名を対象に，第二言語で動画を観る際にいかに字幕が偶発的語彙学習に役立つかを調べた。具体的には(1)英字幕の効果，(2)語彙カバー率の影響，(3)日本語字幕と英字幕の連続的な使用の効果である。参加者を4つの実験群と2つの追加グループに分け，実験一日目に3種類のターゲット語彙事前テストを行った。1週間後の二日目に指定の条件で内容理解を中心に動画を2回視聴してもらい事後テストを行った。これら語彙テストの正誤データに対して，一般化線形混合モデルのロジスティック回帰分析を行った。結果は，どのグループも動画視聴による語彙の学習が見られたが，英字幕，そして日本語字幕と英字幕の連続的な使用による効果は見られなかった。また，字幕の有無に関わらず，95%以上のカバー率がある学習者は有意により多くの語彙を習得できていることが分かった。第二言語での動画視聴に関する教育的示唆についても議論していく。

1 はじめに

近年は動画視聴に関連する第二言語習得研究の報告が数多くある(Wei & Fan, 2022)。ラジオを聴いたり，駅や空港のアナウンスを聞いたりするといった，視覚情報がない聴解活動以上に，動画を観るという活動は昨今の現実世界にあふれている(Graham & Zhang, 2021)といえるかもしれない。この事実は，テレビだけでなくYouTubeやNetflixなどの動画視聴プラットフォームの人気具合から明らかだろう。このように現実世界に即した媒体や方法で外国語を学ぶことは意義深い。

それだけでなく，言語以外の視覚情報に頼らない文章の読解や聴解による学習に比べ，動画視聴による学習は質や量の面でより良くなる場合があるという主張もある(Mayer, 2005)。Richard E. Mayerと彼の共同研究者たちはCognitive Theory of Multimedia Learningを提唱，研究，そして発展させてきた。この理論の要点は，「人間は言葉だけの場合よりも，言葉と画像がともにあることでより良い学びをする(Mayer, 2005, p.1)」ことであり，この主張は人間の心がどのように働くかに関する3つの仮定をもとに理論的に体系化されたものである(Mayer, 2021)。このように，動画視聴による学習は理論の裏付けがあり，効果的かつ効率的であるとされているため，更なる研究は教育現場に対しても有益なものとなるだろう。

動画視聴を用いた第二言語習得の研究では，動画に字幕を加えることで，いかに学習が促進されるかといったことが中心的な話題となっている。字幕付き動画は学習者の内容に関する理解を促進するだけでなく，発音や文法，語彙，そして語用論的知識の獲得にまでも貢献するとされる(Wei & Fan, 2022によるレビュー論文を参照されたい)。その中でも本研究では，外国語環境の学習者が抱える重大な困難の1つが

語彙を学ぶこと(Webb & Nation, 2017), そして日本の文脈で行われた第二言語動画視聴による語彙習得研究が比較的少ないこと(i.e., Dizon & Thanyawatpokin, 2021; Raine, 2013; Rodgers, 2013)を考慮し, 日本人大学生を対象とした字幕付きの第二言語動画視聴による英語語彙学習をテーマとする。字幕というと母語によるものと第二言語によるものが考えられうる。英語で書かれた研究では, 第一言語字幕と第二言語字幕はそれぞれ(L1)subtitlesとcaptionsと呼ばれることが多いが, 本研究では第二言語字幕(つまり英字幕)を中心的に扱いつつも, 第一言語による字幕の是非についても検討するため, 以降の議論では必要がある際, それぞれの字幕を明示する(より詳しい字幕の分類や用語についてはMontero Perez, 2022を参照)。

2 先行研究

2.1 第二言語字幕付きの動画視聴による語彙習得の基礎研究

Montero Perez et al.(2013)は, 第二言語字幕が語彙習得へ肯定的な影響を与えるのか調べた実証研究を使ってメタ分析を行った。メタ分析はテーマに沿って似た研究を集めて, Cohen's *d*やHedges' *g*といった効果量を算出し, 統計的に統合する分析のことである。これにより, ひとつひとつの研究を超えて, ある実験的介入に効果があるのかどうか, 研究の傾向を確認することができる(Borenstein et al., 2021; Plonsky & Oswald, 2015)。Montero Perez et al.(2013)は, 10個の実証研究を統合し, $g = 0.87$ ($p < .001$)という, 中〜大程度の大きさの効果量(Plonsky & Oswald, 2014)を導き出した。つまり, 第二言語字幕付きで第二言語動画を視聴することは, 字幕なしで同じ動画を視聴する場合よりも, 多くの語彙を学ぶことができるという意味である。

しかし, Montero Perez et al.(2013)では考慮されなかった偶発的(付随的)学習の観点を組み込んだ近年の研究では, 第二言語字幕の効果を支持するものとそうでないものが混在している。偶発的学習は, 言語活動の際に学習者の注意が言語の内容に向いているときに, 言語の形式面も併せて学習されることをいう(Hulstijn, 2001)。偶発的学習は, 形式面に全く注意が向かなかったり, 意識がなかったりする暗黙的学習(Ellis, 2009)とは区別される。実験では, 完全に言語形式に関する意図を排除することは難しいが, 後に行われるテストの存在を実験参加者に知らせなかったり, 意図的に参加者の意識を言語の内容面に向けたりすることで, 形式面に関して偶発的学習という構成概念を操作化する(Webb, 2020)。

偶発的学習の環境を作りだし, 第二言語字幕の効果を支持する研究には, 例えばTeng(2022)が挙げられる。彼の研究では, 英語を専攻する中国人大学生1年生82人が, BBCドキュメンタリーを英字幕付きで観るグループと字幕なしで観るグループに分けられた。動画視聴後の4つのターゲット語彙テストの結果を比べると, どの語彙テストにおいても字幕の有無が主効果として検出された。同様にMontero Perez et al.(2018)はドイツ語を母語とし, フランス語を学ぶ大学生を対象とした実験で, 参加者に各2〜4分ほどの異なる動画を3本観させた。ターゲット語彙が視聴した動画に出てきたかどうかを答えるform recognitionと呼ばれるタイプの語彙テスト(Schmitt, 2014)と各ターゲット語彙がどの動画で現れたかを答えるclip associationと呼ばれるテストにおいて, 第二言語字幕の効果を示した。

反対に, 第二言語字幕の効果を発見しなかった研究には, 例えばHao et al.(2022)が挙げられる。Hao et al.(2022)は, 第二言語の熟達度が異なるサンプルそれぞれにおいて, 実験群と統制群を用いて実験を行っているが, ここでの議論は熟達度の高いサンプルに絞る。英字幕付き動画を観るグループと字幕なしで観るグループの参加者は, それぞれの条件で4つのTED Talksシリーズの動画を視聴した。動画視聴後のターゲット語彙テストは与えられた定義に対応する英単語を選ぶform recognitionのテストである。統計分析の結果, 英字幕付きグループは字幕無しグループに比べてテストの結果が有意に悪かった。また, Jelani and Boers(2018)の研究では, マレーシアの高校生を対象に英字幕付きの動画を視聴するグルー

プと字幕なしで動画を観るグループを分け，TED Talksシリーズから1本の動画を2回視聴させた。彼女らの研究で興味深い点は，ターゲット語彙の意味を書くテストで，全体的には確かに英字幕付きグループの方が英字幕なしグループよりも統計的に有意なレベルで得点が良かったが，それはターゲット語彙が書記言語として与えられた場合のみで，ターゲット語彙が音声のみで与えられた場合には得点に違いがなかったのである。

このほかにも偶発的語彙習得に対して第二言語字幕の効果を認める研究が多いものの（e.g., Kang, 2019; Teng, 2018; Wang & Pellicer-Sanchez, 2022），字幕の効果が見られないことを示唆する研究には日本人英語学習者を対象としているものがある。Rodgers（2013）は博論研究として合計5つの研究を行い，そのうち4つは英字幕なしで第二言語の動画を視聴することに関する研究であり，最後の5つ目で英字幕付きの動画視聴に関する研究を実施した。参加者は日本人大学生で，英語の熟達度は初級から中級レベルとされている。10週に渡り，1つ目の研究の参加者は英字幕なしで，5つ目の研究の参加者は英字幕付きでアメリカのドラマの各エピソードを視聴した。語彙テストの結果は，英字幕の有無は偶発的語彙習得に影響がない，とのことだった。同じく未公刊のRaine（2013）はTED Talkを使用し同様の結果を呈している。この原因として，RogersやRaineの研究で使用された動画が語彙や発話速度や訛りなどの音声の観点から，参加者レベルとはマッチしていなかった可能性がある。そのため，より適切なレベルの動画を使用して，日本人英語学習者が英字幕から偶発的語彙学習に関して恩恵を受けられるか，研究の余地がある。

2.2 字幕の効果的な使用に関する発展研究

前述のように，第二言語字幕が必ずしもいつも偶発的語彙習得に寄与するとは限らないことを反映し，近年は第二言語字幕をいかに効果的に扱うかといった観点を持つ研究も見られる。ターゲット語彙や重要な語彙のみを第二言語字幕として提示したり（Montero Perez, et al., 2014），同様に語彙をハイライトして目立たせたり（Hsieh, 2020）することで，視聴者の意識をより語彙に向ける試みもある。ここではいかに学習者の内容理解を促進することで偶発的学習にも結び付けられるかを概観していく。

一つ目は動画の語彙カバー率である。語彙カバー率とはあるディスコースのまとまり（テキスト）において読み手あるいは聞き手が知っている語彙の割合のことであり（van Zeeland & Schmitt, 2013），読解研究の分野で多くの研究がなされてきた。あるテキストの十分な内容理解を得るには，そのテキストに出てくる95%以上，できれば98%の語彙を知っている必要があるとされる（Schmitt et al., 2011）。とはいえ，第二言語動画視聴に関しては，動画の性質上，理解を促進する視覚イメージがあるため，98%以下の語彙カバー率でも十分な理解が得られるのではないかと考えられうる。Durbahn et al.（2020）は，実験参加者一人一人について動画内の既知の単語量を算出し，それを動画の理解度と対照させた。その結果，理解度確認テストの得点が劇的に変化する語彙カバー率の閾値は無かったが，語彙カバー率と動画の理解度に中程度の相関（r = .39）を発見した。下限であった語彙カバー率87%の実験参加者でも62%の理解度を示し，語彙カバー率が88%以降では7割近くの理解度になることから，第二言語動画視聴による十分な理解には，概ね90%以上の語彙カバー率を目指すのが良いと教育的示唆を述べている。しかし，実際にこの90%のカバー率で内容理解のみならず，偶発的に語彙が学習されるかについては検証が必要だ。

また，二つ目の動画の内容理解を促進する手立てとして，母語による字幕（これ以降第一言語字幕と呼ぶ）がある。そもそも第二言語字幕を付けるということは，視聴者がある程度の第二言語読解能力を有していることを想定し（Vanderplank, 2016），動画理解には効率の良い言語処理がかかせない。それに対し，第一言語字幕は文字通り視聴者の第一言語であるため，言語処理に係る認知負荷は明らかに低いと考えられる。第一言語字幕を用いた研究では，確かに字幕がない場合よりも内容理解度テストの結果が良くなることが示されている（Birulés-Muntané & Soto-Faraco, 2016）。しかし，第一言語字幕による語彙

習得の効果は第二言語字幕の場合ほどは見られないことが多い(e.g., Frumuselu et al., 2015; Peters, 2019)。これは, 第一言語による意味とそれに対応する第二言語の語彙形式を音声から対応させることが難しいためかもしれない。特に日本語と英語の関係のように, 第一言語と第二言語の言語距離が遠い場合には, 一般的に言語習得は第一言語と第二言語の言語距離が近い場合よりは難しいとされ(Wesche & Paribakht, 2010), 統語論的な違いや使用される語彙の隔たりのために, 言語間の対応が見分けづらいのであろう。

そこで, 効果的な第一言語字幕の使用法として, 第一言語字幕付きで動画を視聴した後に第二言語字幕付きで同じ動画を観るといった連続的な字幕の使い方が考えられる。第一言語字幕付きで第二言語の動画を視聴する場合は, 提示される書記言語が母語のため言語処理が浅くなってしまう可能性が否めないが, それでもある程度の動画理解が心的表象として作成されたならば(Kintsch, 1998), その知識をもとに2回目の動画視聴における理解が飛躍的に上がり, 偶発的学習の土台となったり, あるいは2回目の動画視聴ではより言語形式面への注意が向いたりするかもしれない。Wi and Boers(2023)は, カナダの大学で学ぶ中国人英語学習者を対象に第一言語字幕と第二言語字幕の連続的な使用を調べた研究である。使用された動画はTED Talksであり, 実験群は, 1回目の視聴を第一言語字幕付きで, 2回目を第二言語字幕付きで, 3回目の視聴を字幕なしで視聴し, 統制群は3回とも第二言語字幕付きで動画を視聴した。混合ロジスティック回帰分析の結果, ターゲット語彙の意味を書く語彙テストで, グループ間の違いは見られなかった。しかし, この結果を一般化してしまうのは早急である。というのも, Wi and Boers(2023)が認めるようにサンプルサイズがそもそも小さく(実験群n = 11, 対照群 n = 9), 且つターゲット語彙が11個のみであったためだ。また, 実験デザインの観点からも, 彼らの実験参加者は合計3回の動画視聴ができたため, 特に統制群における語彙の伸びの大きさは回数による影響が否めない。そのため, 第一言語字幕と第二言語字幕の連続的な使用については引き続き研究する価値があるといえる。

3 研究設問

先行研究のセクションで概観したように, 偶発的語彙学習に対する第二言語字幕付き動画視聴の影響を調べた基礎研究は日本では比較的稀有である。語彙カバー率の観点から適切なレベル(i.e., 概ね90%以上)の第二言語動画で日本人英語学習者が内容理解だけでなく, 偶発的に語彙を学べるとすると, 教育現場に応用しうる示唆を得ることができる。また, 実際の語学学習を考えると, 字幕のみを用いて動画を視聴したり, その逆に字幕を全く使わないで視聴したりする状況はあまりないだろう。字幕をどのように効果的に組み合わせ得るかという視点が第二言語動画視聴の応用研究に必要である。更に, 使用する動画の語彙カバー率も学習教材の選定にあたって非常に重要な視点である。そのため, 本研究では以下3つの研究設問を立てて, 字幕効果を探っていく。

研究設問1 英語を外国語として学ぶ日本人大学生(及び大学院生)は, 英字幕付きで英語の動画を視聴するほうが, 字幕なしで同じ動画を視聴するよりも偶発的に学ばれる語彙量が多いか。

研究設問2 英語の動画を視聴する際, 学ばれうる語彙量と動画の語彙カバー率との間に関連があるか。

研究設問3 英語を外国語として学ぶ日本人大学生(及び大学院生)は, 英語の動画を1回目は日本語字幕付きで2回目は英字幕付きで視聴するほうが, 同じ動画を1回目は字幕無しで2回目は英字幕付きで視聴するよりも偶発的に学ばれる語彙量が多いか。

4 本研究について

4.1 参加者

データ分析の対象となる本研究の参加者は，研究を紹介するフライヤーを通じて参加を募った合計148人の日本語を母語とする大学生，大学院生である。募集要項では参加条件を日本語が母語あるいは母語に準ずるレベルで，基本的に日常では日本語を使用すること，英語圏での海外経験が通算6か月未満であること，英語力がおおよそ2年以内の語学試験の結果でCEFRレベルB1からＢ２の範囲（英検：２級か準1級，TOEIC L&R: 550-935, TOEFL iBT: 42-94, IELTS: 4.0-6.5）であることを明示した。このレベル帯に絞った理由は，動画で十分な理解が得られることを担保するためと，おおよそ3000語レベルの語彙知識を有し，且つ英字幕を十分に読解で処理できる能力があるような参加者を募るためである。

4.2 使用動画とターゲット語彙

使用する動画は，理解しやすいように，動画内で使用される語彙に対応するイメージが多いことが予想されるドキュメンタリーかアニメーションから選ぶことを基本とし，且つ高頻度3000語レベルで90%以上の語彙カバー率があること（Durbahn et al., 2020）を条件とした。その結果，TED-Edシリーズから“The Evolution of the Book（https://www.youtube.com/watch?v=_YqYtdPUis4）”を選出した。TED-Edシリーズは科学的，学術的に興味深い話題をアニメーション形式で解説する動画だ。動画内容に関係のない最初と最後の部分を除くと，長さは3分59秒であり，総語数が618語のため，約155.1WPM（= Words Per Minute）の速さであった。このWPMでは速すぎる可能性を考慮し，ナレーションの発音が不明瞭にならない0.9倍速で動画を再生することに決定した。この結果，動画の速さは約139.6WPMとなった。また，動画で用いられている語彙カバー率の算出については，lextutorサイト上のCompleat Web VP v.2.6（Cobb, n.a.）を使用し，結果は以下の表1のとおりである。

■表1: 使用動画の語彙カバー率

頻度レベル	異なり語数（%）	総語数（%）	語彙カバー率（%）
1k	205 (65.7)	485 (78.5)	78.5
2k	60 (19.2)	76 (12.3)	90.8
3k	23 (7.4)	26 (4.2)	95.0
4k	8 (2.6)	12 (1.9)	96.9
5k	5 (1.6)	5 (0.8)	97.7
6k	2 (0.6)	2 (0.3)	98.0
7k	2 (0.6)	4 (0.6)	98.6
8k	2 (0.6)	3 (0.5)	99.1
10k	1 (0.3)	1 (0.2)	99.3
12k	2 (0.6)	2 (0.3)	99.6
13k	1 (0.3)	1 (0.2)	99.8
リスト外	1 (0.3)	1 (0.2)	100.0
合計	312 (100)	618 (100)	≈100.0

注．頻度レベル列におけるkは1000を意味する。

表1の頻度レベルの列でそれぞれ3kと2kの行をみると最頻3000語で動画内の語彙が95%カバーされ，2000語でも90.8%カバーされるため，2000から3000語レベルの語彙知識を持つ学習者にとっては，動画の十分な理解が得られると考えられる。このため，語彙レベルの観点から本動画は適切なレベルであると言える（Durbahn et al., 2020）。ターゲット語彙は表1を参考にしつつ，低頻度語や高頻度語であってもなるべく参加者が知らない可能性のある英単語を18語，品詞を限定することなく抽出した。このように選ばれたターゲット語彙は，alloy, binding, boom, carve, durable, dye, elite, flexibility, in bulk, mixture, mold, on-the-go, parchment, pop up（動詞），profound, soot, stack, walnutである。

4.3 テスト

今回使用するテストは，動画の理解を測る内容確認テスト，頻度3000語レベルまでの語彙知識量を測る語彙レベルテスト（VLT），そして動画視聴によるターゲット語彙の学習効果を測る語彙テスト3種類である。語彙テストは，形式認識テスト，意味認識テスト，そして意味記述テストの3種類であるが，複数のテストを行う理由は，語彙習得には段階があり，1種類のみのテストでは学習者の語彙習得をきめの細かいレベルで測定することができない可能性があるためである（Webb, 2007）。

また，内容確認テストは1回目の動画視聴後にのみ行われたが，ターゲット語彙テスト3種類は事前テスト，事後テスト，そして遅延事後テストとして実施した。これらのターゲット語彙テストはすべて同じものであるが，テスト内の項目は参加者ごと，且つテストを受ける回ごとにランダマイズされるようになっている。ターゲット語彙テストが難しくて参加者のモチベーションを下げることにならないように，分析には含めない最頻2000語レベルまでのフィラー語も10単語それぞれのテストに含めた。

4.3.1 内容確認テスト

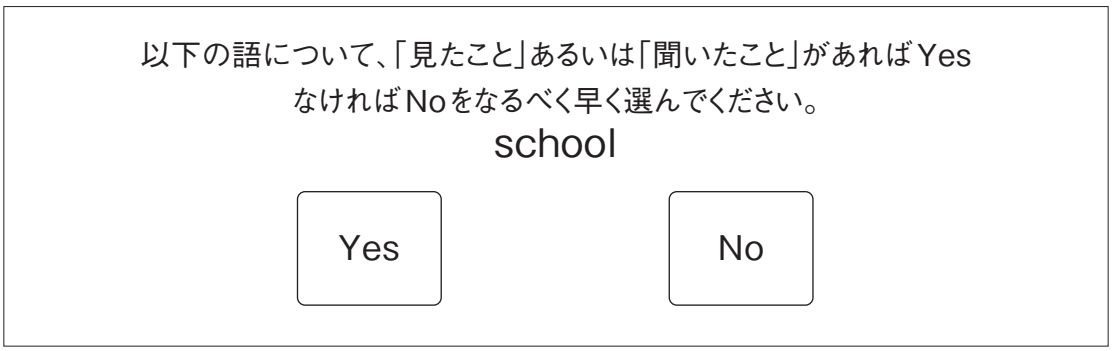
内容確認テストを用いる主たる理由は，参加者の意識を内容理解に向けさせ，動画理解に集中させるためである。実験の時間的制約があり，記述式の問題ではなく，全10問の多肢選択式問題をGoogle Formを用いて作成した。このテストで扱う内容は，ターゲット語彙テストにおいてヒントとはならないように，ターゲット語彙が現れる文とは関連がないように配慮した。また，内容確認テストは指示文や選択肢を含め，すべて日本語で作成された。なお，本実験での関心は内容理解というより語彙習得にあることに注意されたい。

4.3.2 語彙レベルテスト

本実験で用いる語彙レベルテストには，Webb et al.（2017）によるアップデートされた版の3000語レベルまでを使用する。各レベルは30問ずつあり，定義にあてはまる英単語を選ぶことが求められる。26語以上の正答数でそのレベルをマスターしていると考える。このテストを用いて本実験参加者が動画視聴をするのに十分なレベルの語彙を有しているかを確認する。語彙レベルテストの結果，1000, 2000, 3000のすべての頻度レベルで26語以上正解している場合には，動画内の全単語の「95%以上」をカバーしているものとし，1000と2000のレベルのみ26語以上正解している場合は「90%以上（95%未満）」，そして前者2パターンのいずれでもない場合は「90%未満」とみなすこととした。

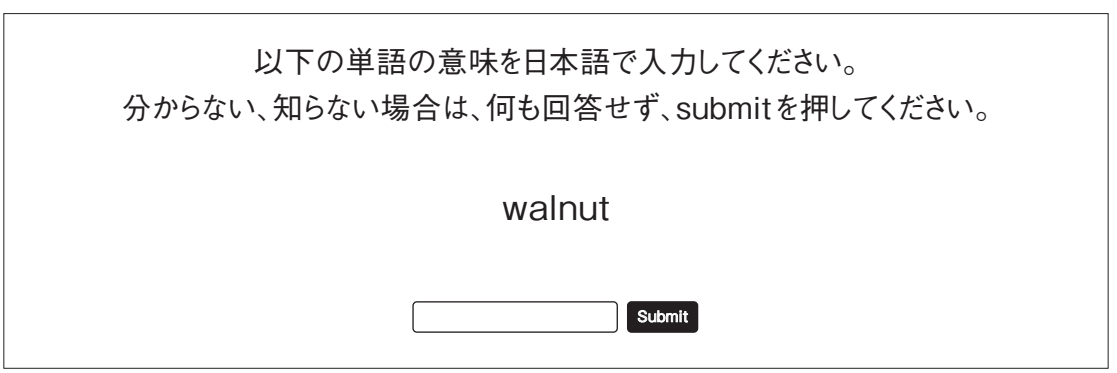
4.3.3 形式認識テスト

形式認識テストは，提示された単語を見たこと，あるいは聞いたことがあるかを問うものである。意味に関する知識を持っている必要はないので，その意味で今回使用する3種類のターゲット語彙テストのうち一番簡単なレベルのテストであると言える。三種類のターゲット語彙テストの中で最初に実施された。見聞きしたことがあれば「はい（yes）」を選択し，そうでなければ「いいえ（no）」を選択するため，前者を1，後者を0とする二値データとして自動採点した。クロンバックのαは事前，事後，遅延事後テストでそれぞれ，α = .65, .75, .84であった。以下は形式認識テストの例であるが，提示されている語は例題として出されたものである。



4.3.4 意味記述テスト

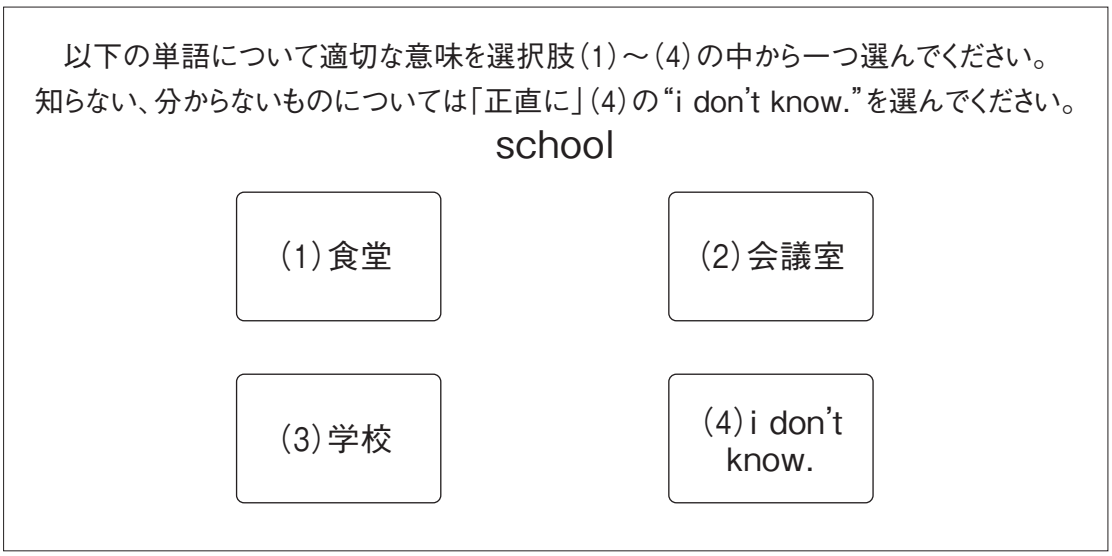
意味記述テストは，提示された単語の日本語訳，あるいは説明を書くものである。単語の形式をもとに，自力でメンタルレキシコン内の意味にアクセスし，取得する必要があるため，今回の語彙テストの中では一番難しいと考えられる（Laufer & Goldstein, 2004）。なお，本テストは後述の意味認識テストよりも先に行われたが，それは意味認識テストを先に行った場合，選択肢が意味記述テストのヒントになってしまう恐れがあるからである。意味記述テストの採点については，筆者及び，筆者同様に第二言語習得論や言語テスト理論を専攻する修士課程の学生の力を借りて採点した。採点基準は，ターゲット単語を動画内で使用される意味で回答できれていれば1，そうでなければ0として二値データになるよう採点した。例えば，今回の動画ではpop upは動詞で「現れる」という意味で使用されているため，同じ意味の「出現する」であれば正解とし，「跳びあがる」は不正解とした。採点手順については，筆者と学生で採点基準や採点方法に関する摺り合わせの後に，全データの30%にあたる2940項目の回答を各々採点した。採点者間信頼性を示すカッパ係数は0.78となり，Plonsky and Derrick（2016）のメタ分析で示されるよりは幾分値が低いものの十分な程度と考えられるため，残りの回答データはすべて筆者が採点した。クロンバックのαは事前，事後，遅延事後テストでそれぞれ，α = .51, .62, .67であった。以下は意味記述テストの例であり，提示されている語も本題として出されたものである。



4.3.5 意味認識テスト

意味認識テストは，提示された単語の正しい意味を三択あるいは「知らない（I don't know）」から選ぶものである。提示された選択肢の中から該当する意味を選ぶため，語の形式と意味のつながりが強固である必要はない。そのため，今回の3種類の語彙テストの中では2番目の難しさであるが，前述の理由で最後に行われた。本テストも正しい意味が選べていれば1，不正解あるいは「知らない」という選択肢を選んだ場合には0とする二値データとして自動採点した。クロンバックのαは事前，事後，遅延事後テストでそれぞれ，α = .57, .68, .70であった。以下は意味認識テストの例であるが，提示されている語は例題として出されたものである。例題の指示文にあるように，推測の影響を出来る限り除くため，知らない単語が出

てきた場合，参加者は4つ目の選択肢である「知らない（I don’t know）」を選ぶことが求められた。



注．上記画像ではI don’t know. とすべきところが，実験作成時のミスで小文字のiになってしまっている。

4.4 実験の流れ

実験の全工程は，2022年の9月から開始し翌年3月まで行われた。研究を紹介するフライヤーを通じて，実験参加希望者にはGoogle Formでコンタクトに必要な個人情報や参加希望日時を記入してもらった。その後参加者に実験一日目用のzoomリンク，あるいはzoom録画を送り，実験全体の流れや実験参加は任意であること，いつでも参加辞退が可能であることを説明した。実験参加希望者には，実験参加同意書の記入と提出を求めた。実験一日目は，VLTや事前語彙テスト3種類など，Gorilla Experiment Builder（Anwyl-Irvine et al., 2020）を使用して作成した種々のテストを継続して受験してもらった。なお，実験参加者は申し込みの確定順に全6グループのうちいずれかのグループにランダムに割り振られていった。各グループの詳細については以下表2のとおりである。Exp1とCon1は1つ目の研究設問に答えるための実験群と統制群であり，Exp2とCon2は3つ目の研究設問に答えるための実験群と統制群である。なお，4グループ以外にも追加で2グループを含めている。英字幕付きで2回動画を観させ，2回目は意識を未知語に向けさせたIntグループを含めた理由については後述する。動画を一度のみ，しかもそれは字幕だけでなく音声もつけなかったBaseグループを含めたのは，今回使用した動画がアニメーションだけでもどれだけ理解しやすいのかを調べるためと，語彙テスト結果のベースラインにするためである。

■表2: 実験グループの詳細

グループ名	人数（*n*）	1回目の視聴	2回目の視聴	補足
Exp1	25	英字幕付き	英字幕付き	研究設問1の実験群
Con1	25	字幕なし	字幕なし	研究設問1の統制群
Exp2	25	日本語字幕付き	英字幕付き	研究設問3の実験群
Con2	24	字幕なし	英字幕付き	研究設問3の統制群
Base	25	音声，字幕なし	2回目の視聴なし	ベースライングループ
Int	24	英字幕付き	英字幕付き	2回目の視聴は意図的学習

実験一日目のおおよそ1週間後，参加者には再度zoomに集まってもらう，あるいは特別な事情の際には実験の録画を共有するというかたちで，実験二日目を実施した。参加者には動画を観てもらうことを伝え，その際に動画視聴後に内容確認テストがあるので，内容理解に努めてもらうこと，それと同時に動画自体を楽しんでもらうよう伝えた。偶発的語彙学習環境を創り出すために，語彙テストがあることは伝えなかった。グループに応じた方法で動画を1回視聴してもらったのちに内容確認テストを実施した。その後，内容の再確認のため動画を再度観てもらうことを伝え，グループに応じた方法で動画を2回目の視聴を実施した。2回目の動画視聴後には，事前テストで使用したのと同じ語彙テストを事後テストとして実施した。更に1週間後の三日目に，参加者にアンケートと合わせて，語彙テストを遅延事後テストとしてメールで送信した。以上の実験工程は表3にまとめられている。

■表3: 実験の流れと内容

実験日	内容
一日目	①実験説明 ②VLTとターゲット語彙事前テスト
二日目	①動画視聴1回目 ②内容確認テスト ③動画視聴2回目 ④ターゲット語彙事後テスト
三日目	①アンケートとターゲット語彙遅延事後テスト

4.5 分析

分析には無料統計ソフトR（R core team, 2023）と関連パッケージを使用した。具体的には，前述の採点者間信頼性の計算にirrパッケージ（Gamer & Lemon, 2019），内的一貫性指標クロンバックのα算出にpsychパッケージ（Revelle, 2023），ルビーン検定にcarパッケージ（Fox & Weisberg, 2019），混合効果モデルの作成，検証にlme4パッケージ（version 1.1-31; Bates et al., 2015），データの処理統合にdplyrパッケージ（Wickham et al., 2023）である。

研究設問に答えるためのターゲット語彙テストのデータは事前テストと事後テストに限る。これは本実験では各テスト間に一週間の空きがあり，遅延事後テストまで取り組んでもらえないことが多々あったためである。分析対象となるデータは，参加者一人一人のターゲット語彙テストにつき，事前テストで正答した項目を事後テストの結果から除いたものとする。例えば，参加者Aさんが事前テストで単語Xを正解しているならば，同じAさんの事後テストの結果からは単語Xのデータを除くということである。こうすることで，もともと正解し習得されていたと考えられる単語を除き，学ばれる余地のある単語のみに絞った分析及び考察が可能になる。分析では，それぞれ研究設問の対象となるグループ（Exp1 vs. Con1あるいはExp2 vs. Con2）の各語彙テスト（前述のように，各参加者につき事前テストで正解だった単語を事後テストから除いたデータセット）に対して一般化線形混合モデルを構築する。目的変数である語彙テストデータは正解か不正解か，つまり1か0かの二値データであるため，二項分布となり，構築される一般化線形混合モデルはロジスティック回帰モデルとした（Brown, 2021）。

モデルに含めた固定効果（本稿で興味のある変数）は，グループ間の違いと語彙カバー率，そしてそれらの交互作用である。これに対してランダム効果（直接的には興味のない変数；詳しくはGarcia, 2021を参照）はターゲット語彙と参加者とした。モデルはCunnings（2012）に見られるように，1つずつ固定効果を増やしてシンプルなモデルから複雑なものを構築していく流れをとり，カイ二乗検定の結果ならびに赤

池情報量基準(AIC)とベイズ情報量基準(BIC)の値などを総合的に確認することでベストフィットなモデルを選択した。なお, 紙面上の都合から, 結果で提示するのは最終的なモデルに限る。

5 結果と考察

5.1 記述統計量

5.1.1. 内容確認テスト

内容確認テストの記述統計量は表4のとおりである。グループに等分散性があるか確認するルビーン検定の結果, それぞれExp1とCon1のグループ間, Exp2とCon2のグループ間に等分散の仮定をしてよさそうである(それぞれp = .950, .437で対立仮説を棄却)。そのため, 等分散を仮定した対応のないt検定を行い, 両グループ間ともに内容確認テストの平均値の差に有意な違いがある結果となった(Exp1とCon1のグループ間は, t = 2.102, df = 48, p = .041, Exp2とCon2のグループ間は, t = 3.140, df = 47, p = .003)。そのため, 英字幕, あるいは日本語字幕付きで第二言語動画を観るほうが, 字幕がない場合に比べてより良い内容理解に貢献するといえる。

Baseグループの結果から, グループ平均点は他のどのグループよりも低いものの, 今回使用したアニメーション動画は映像だけでも内容が理解しやすいものであったことがうかがえる。

■表4: 内容確認テストの記述統計量(満点10点)

グループ	*n*	*M*	*SD*
Exp1	25	7.72	1.25
Con1	25	6.92	1.38
Exp2	25	8.28	1.22
Con2	24	7.17	1.21
Base	25	6.56	1.24
Int	24	7.00	1.76

5.1.2. ターゲット語彙テスト

形式認識テスト, 意味認識テスト, そして意味記述テストの得点の伸びに関する記述統計量は表5のとおりである。前述のとおり, 表5のデータは, 各参加者につき, 事前テストで正答した項目を事後テストから除いたものであるため, 平均値(M)が指すのは事前テストから事後テストでの得点の伸びである。どのグループ, どの種類のターゲット語彙テストに関しても得点の伸びが見受けられる。ただし, Baseグループは動画を音声も字幕もなしで視聴しているため, 語彙を習得する余地のないはずであるが, 実際にはどのテストにおいても得点の伸びが確認される。このうち形式認識テストにおいては, 事前テストでターゲット語彙に触れているため, 事後テストでもそれらが見聞きしたことのある単語として認識され, 得点が上がったのだろう。意味記述テストの結果は他のグループと似た結果になっているため, 事前テストから事後テストまでの一週間の期間に実験以外で学習が生じたかもしれない。Baseグループのテスト結果から, 今回の学習の伸びは必ずしもすべて実験によるものではない可能性があることを注意点として挙げておく。

■表5: ターゲット語彙事前テストから事後テストの伸びに関する記述統計量

グループ	形式認識テスト			意味認識テスト			意味記述テスト		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Exp1	25	3.04	1.49	25	2.00	1.68	25	2.20	1.29
Con1	25	2.52	1.90	25	2.80	1.66	25	2.08	1.61
Exp2	24	3.21	1.72	25	2.56	2.35	25	1.92	1.55
Con2	24	3.00	1.82	24	2.79	1.79	24	1.42	1.25
Base	25	2.24	2.01	25	1.36	1.52	25	2.04	1.90
Int	23	3.61	2.08	23	2.43	1.95	24	1.96	1.73

注. 実験ソフトの問題でデータがとれていなかった参加者のテストデータは除いたため, 参加者が少なくなっている場合がある。

5.2 研究設問1&2

Exp1とCon1グループのターゲット語彙テストデータに対して一般化線形混合モデルを当てはめた結果は表6のとおりである。採用されたモデルの固定効果はグループの違いとカバー率を含むが, これら2つの変数間の交互作用はデータ不足のためモデル化できなかった。

■表6: Exp1とCon1グループに対する一般化線形混合モデル分析の結果

	固定効果	オッズ比	*SE*	95%信頼区間	*z*	*p*
形式認識テスト	(Intercept)	0.59	0.30	[0.22, 1.62]	-1.02	.308
	グループ[Exp1]	2.64	1.05	[1.21, 5.77]	2.44	.015
	カバー率[>90%]	1.61	0.84	[0.58, 4.46]	0.92	.355
	カバー率[>95%]	2.51	1.33	[0.89, 7.08]	1.74	.082
	ランダム効果	分散	*SD*			
	参加者	0.80	0.89			
	ターゲット語彙	0.72	0.85			
意味認識テスト	固定効果	オッズ比	*SE*	95%信頼区間	*z*	*p*
	(Intercept)	0.39	0.18	[0.16, 0.95]	-2.07	.039
	グループ[Exp1]	0.74	0.26	[0.37, 1.49]	-0.84	.403
	カバー率[>90%]	1.06	0.52	[0.41, 2.77]	0.13	.898
	カバー率[>95%]	1.92	0.90	[0.77, 4.81]	1.40	.163
	ランダム効果	分散	*SD*			
	参加者	0.77	0.88			
	ターゲット語彙	0.55	0.74			
意味記述テスト	固定効果	オッズ比	*SE*	95%信頼区間	*z*	*p*
	(Intercept)	0.06	0.03	[0.02, 0.17]	-5.13	< .001
	グループ[Exp1]	1.33	0.40	[0.74, 2.40]	0.95	.344
	カバー率[>90%]	1.26	0.55	[0.54, 2.95]	0.54	.587
	カバー率[>95%]	3.69	1.55	[1.62, 8.40]	3.11	.002
	ランダム効果	分散	*SD*			
	参加者	0.40	0.63			
	ターゲット語彙	2.56	1.60			

英字幕が有り無しというグループの違いが有意な影響を持ったのは形式認識テストのみにおいてであった。英字幕があるグループのほうが事後テストで約2.6倍ターゲット語彙の形式を認識できることが分かる(p = .015)。また, カバー率については, 95%以上のカバー率が意味記述テストにおいてのみ有意に肯定的な影響を持つことが分かる。つまり, 動画内の単語の95%以上を知っている参加者は, 約3.7倍ターゲット語彙の意味を筆記テストで正しく産出することができるということである(p = .002)。まとめると, 英字幕があることで結果が良くなるのは形式認識テストのみで, 95%以上のカバー率があると, 英字幕の有無に関わらず, 意味記述テストの結果がより良いということである。

形式認識テストにおいて英字幕が有意に好影響を持ったのは, インプットを聴覚だけでなく視覚も合わせて受けることで, 未知語が際立つのではないかと考えられる。しかし, 意味認識テスト及び意味記述テストのモデルから, 英字幕はターゲット語彙を未知語として際立たせるだけで, 英字幕が無い場合よりもターゲット語彙の意味を推測しやすくさせることはなさそうである。

この結果は, 偶発的語彙習得に対する第二言語字幕の効果を発見しなかった先行研究(e.g., Raine, 2013; Rodgers, 2013)を支持するものである。第二言語字幕の効果がなかった理由としてまず考えられるのは, 本実験の動画が言語音声面で難しかったわけではないことだ。第二言語字幕は特にリスニングの苦手な学習者に対してセグメンテーション(明確な区切りのない音声言語を語彙単位で区切っていくこと;Rost, 2016)の手助けとなるとされるが(Charles & Trenkic, 2015), Ted-Edのナレーターは訛りなく非常に分かりやすい発音で話しており, 話の速さも約137.1WPMであったため, 英字幕のないグループの参加者にとってもセグメンテーションがしやすく, それによりリスニング面で大きな負担がなかったのかもしれない。つまり, 本実験の対照群の結果は, 仮に本実験の動画の発話速度が速かったり, 訛りが強かったりするなど音声的に難しかった場合と比べて, より良いものであった可能性がある。しかし, この説明は内容確認テストの結果と照らし合わせるといささか相容れない。というのも, 実際に英字幕グループは内容確認テストの結果が英字幕なしのグループに比べて有意なレベルで良いため, 英字幕の恩恵が内容理解にあったことを示唆する。それはつまり, 英字幕があったことにより英字幕グループの方がセグメンテーションがより良くできた可能性がある。そのため, 動画の音声面の易しさにより英字幕が偶発的語彙学習を促進しなかったという説明は必ずしも適切でないかもしれない。

そこで2つ目に考えられる理由としては, 英字幕の有無に関わらず, ターゲット語彙に向けられた注意の量, つまりはターゲット語彙の言語処理量が両グループで同じくらいであった可能性だ。本実験では, あらかじめ参加者に動画視聴の際は内容理解に努めて動画を楽しむよう指示を出しているため, その差は参加者ごとに違えど「英単語を学ぼう」という意識はあまりなかったはずである。すると, 実験群と対照群のどちらにおいても, 有限である注意(Mayer, 2021)がほとんど内容理解に使われ, ナレーションであれ英字幕であれ言語的側面(未知語の意味)に向けられる注意の総量がどちらのグループでも違いはほとんどなかったのかもしれない。

次に研究設問の2つ目となるカバー率と語彙習得の関係性について議論をしていく。本実験ではVLTの結果をもとに, 各参加者が動画内に出てくる単語をどれだけ知っているかを, 95%以上, 90%以上(95%未満), 90%未満の3レベルで分けた。このカバー率の情報を統計モデルに組み込んだ結果, 意味記述テストにおいて有意に肯定的な影響が確認された。これはつまり, 動画内で出てくる単語の95%以上を知っている場合, 未知語の意味を正しく推測し, それを自力で再生できる度合いが大きくなるということである。動画内で出てくる単語の多くを知っているということは必然的に未知語の割合が少ないため, 限られた注意を限られた数の未知語に向け, イメージを含む周りの文脈から正しく未知語の意味を推測できることを示唆する。逆に言えば, カバー率が95%未満の場合は, 多くの未知語が出てきて内容理解に多くの注意量を割いてしまうため, 未知語の推測どころではなくなってしまっている可能性がある。今回のモデルでは残念ながらこのカバー率とグループの違いの交互作用をそもそも含むことができなかったため, 95%以上のカバー率があり, 加えて英字幕があればより多くの語彙量を習得できたかについては今後の課題である。

5.3 研究設問3

Exp2とCon2グループのターゲット語彙テストデータに対して一般化線形混合モデルを当てはめた結果は表7のとおりである。1回目の視聴で日本語字幕あり，なしというグループの違いはどの種類のターゲット語彙テストにおいても有意な結果とはならなかった。つまり，日本語字幕と英字幕を連続的に使用した場合と，1回目に字幕がなしで2回目に英字幕付きで動画を視聴した場合に，字幕の違いは偶発的語彙習得に影響を与えなかったことを意味する。また，カバー率については，研究設問1&2のモデルと同様に，95%以上のカバー率が意味記述テストにおいてのみ有意に肯定的な影響を持つことが分かる。動画内の単語の95%以上を知っている参加者は，約4.3倍ターゲット語彙の意味を筆記テストで正しく産出することができるということである（p = .001）。グループの違いとカバー率の2つの変数間の交互作用はデータ不足のためモデル化できなかった。

■表7: Exp2とCon2グループに対する一般化線形混合モデル分析の結果

	固定効果	オッズ比	*SE*	95%信頼区間	*z*	*p*
形式認識テスト	（Intercept）	1.17	0.63	[0.41, 3.34]	0.29	.775
	グループ [Exp2]	0.78	0.30	[0.37, 1.66]	-0.64	.522
	カバー率 [>90%]	2.14	1.10	[0.79, 5.84]	1.49	.135
	カバー率 [>95%]	2.60	1.38	[0.92, 7.36]	1.80	.072
	ランダム効果	分散	*SD*			
	参加者	0.67	0.82			
	ターゲット語彙	0.69	0.83			
意味認識テスト	固定効果	オッズ比	*SE*	95%信頼区間	*z*	*p*
	（Intercept）	0.38	0.17	[0.16, 0.93]	-2.12	.034
	グループ [Exp2]	0.83	0.27	[0.44, 1.58]	-0.56	.575
	カバー率 [>90%]	0.98	0.44	[0.41, 2.36]	-0.04	.972
	カバー率 [>95%]	1.50	0.69	[0.61, 3.68]	0.88	.377
	ランダム効果	分散	*SD*			
	参加者	0.58	0.76			
	ターゲット語彙	0.53	0.73			
意味記述テスト	固定効果	オッズ比	*SE*	95%信頼区間	*z*	*p*
	（Intercept）	0.04	0.02	[0.01, 0.12]	-5.76	< .001
	グループ [Exp2]	1.37	0.39	[0.78, 2.39]	1.10	.272
	カバー率 [>90%]	1.81	0.79	[0.77, 4.27]	1.35	.176
	カバー率 [>95%]	4.32	1.93	[1.80, 10.37]	3.28	.001
	ランダム効果	分散	*SD*			
	参加者	0.14	0.37			
	ターゲット語彙	2.22	1.49			

Exp2とCon2のグループの結果も，Exp1とCon1グループの結果同様，偶発的語彙習得に影響を与えるのは，字幕の有無というより第二言語での動画視聴そのもののようである。この結果は同じく第一言語と第二言語の字幕の連続的な使用について研究したWi and Boers（2023）と似た結果である。しかし，本実験でExp2とCon2で字幕効果がなかった理由としては，Wi and Boers（2023）で述べられているようなサンプルサイズの問題ではなく，研究設問1で述べたような注意量が関係しているかもしれない。

あるいは，Exp2とCon2のテスト結果に違いがないのは，日本語字幕が内容理解には好影響を持つが，

英語字幕と併用されたとしても偶発的語彙学習には影響がない可能性もあり得る。日本語字幕と英語字幕の連続使用を2つ目の研究設問として調査したのは，日本語字幕が内容理解を促進し，英語字幕付きで動画を2回目の視聴をしたときに，未知語の形式と意味を結びつけやすくなるとの予想のためである。実際に実験群のほうが統計的に有意なレベルで内容理解が優れているが，動画視聴による偶発的語彙学習に内容理解の度合いはあまり関係せず，偶発的語彙学習には，後に触れるカバー率のような別の要素が大きな影響を持つのかもしれない。

最後にもう一つの可能性を挙げる。本実験の記述統計量を確認すると（表5参照），Exp1, Exp2, Con1, Con2のすべてのグループが似通った結果であるのが分かる。このことは，本実験で偶発的語彙習得の量にグループ間の違いが見つけられなかったもう一つの可能性を指し示すものである。それは視聴回数である。視聴回数に絡めながら本実験の流れを振り替えると，どのグループも指定の条件で1回目の視聴を終えた後には内容確認テストを受けている。内容確認テストの結果は，それぞれの実験で字幕付きのグループのほうが統計的により良いものであり，それは1回目の視聴時点で字幕の条件に応じて参加者に何らかの質的な違いがあったことを示唆するものである。しかし，ターゲット語彙テストについては1回目の動画視聴終了時点で実施したのではなく，もう一度参加者に指定の条件で動画を視聴してもらった後に実施した。2回目の視聴チャンスがあったことで，2回目に字幕がなかったグループ（Con1とCon2）の参加者は2回目の視聴でより集中して動画を視聴し，反対に1回目に字幕があったグループの参加者は既に十分な内容理解が得られていると考え，2回目の視聴がおろそかになってしまった可能性が挙げられる。結果的に実験群と対照群の学習が似通ったのかもしれない。

カバー率については，2つ目の研究設問と同様に，動画内で出てくる単語の95%以上を知っている場合，意味記述テストの結果がより良くなることが確認された。結果が一つ目の研究設問と似ているため，音声面でそれほど難しくない動画を視聴することによる偶発的語彙学習で大切なことは，字幕の有無というより，動画内で出てくる語彙を95%以上知っていることかもしれない。

6 まとめと課題

本研究では，第二言語動画を字幕付きで観ることが偶発的語彙習得を促進するかといったテーマで研究を進めてきた。最後に各研究設問をまとめつつ，教育的意義や今後の研究の課題をまとめていく。

まず1つ目の研究設問では，日本人大学生が，英字幕付きで英語の動画を視聴する際に，字幕なしで同じ動画を視聴する場合よりも偶発的に学ばれる語彙量が多いかを問うた。本実験の結果は予想に反して，偶発的に学ばれる語彙量は英字幕の有無に関係しなかった。英字幕の有無が影響を与えなかった理由としては，語彙に向けられる注意量が実験群でも対照群でも同じだった可能性や，使用された動画の音声的な難しさの観点からも議論した。ただし，本実験の結果は，英字幕の効果を全否定するものではない。実際に本実験では実験群の内容理解は有意に対照群よりも良かったし，英字幕の効果は偶発的語彙習得だけでなく，発音や語用論的知識など多角的に研究されてきている（Wei & Fan, 2022）。そのため，授業内あるいは自主学習において第二言語で動画を視聴する場合には，細部に至る内容理解を得るために第二言語字幕を付けることを積極的に選択肢に入れてみるのも良いだろう。しかし，英字幕を付けると英文を読むだけになり，動画の音声面に注意がいかなくなってしまうのではないか，という危惧もある。確かに，初級レベルの学習者によってはその可能性が否めないが，中級レベル以上の学習者であれば第二言語字幕付きの動画を観ている際，リスニングにも注意が向いていることを示す研究もある（Charles & Trenkic, 2015）。そして動画を選ぶ際に重要なことは，面白そうな動画をやみくもに選ぶのではなく，内容理解に関しては動画内に出てくる語彙の90%以上（Durbahn et al., 2020）を，そして偶発的な語彙の習得に関しては95%以上を学習者が知っているような動画を選択することであると言える。

また，第二言語字幕付きの動画を使って偶発的に単語を学ばせるのではなく，意図的に学ばせるのはど

うかといった考えもあるだろう。本実験で含めた6つ目のIntグループは，2回目の視聴で未知語があればそれに注意を向けるよう指示されたグループであるが，記述統計量を見る限り（5.1.2参照），結果は形式認識テストを除いて偶発的学習グループ群の結果とあまり違いがなさそうである。テストで測定された語彙量に違いはなさそうだが，偶発的語彙学習の良さは明示的に教えるのが困難な単語の意味が学習されたり，文脈と合わせた使用法や意味が獲得されたりすることだとされている（Schmitt, 2010）。もし教室内外での語彙指導及び学習が意図的学習に偏るものであれば，字幕を組み合わせた第二言語動画視聴による偶発的学習の比重を増やすのも良いかもしれない。もちろん，偶発的学習と意図的学習は相互排他的なものではなく，どちらの学習タイプもうまく使われていく必要がある。

3つ目の研究設問では，日本語字幕と英字幕の連続的な使用を，1回目は字幕が無い状態から2回目は英字幕付きで視聴する場合と対照させて，偶発的に学ばれる語彙量に違いがあるかを検証した。残念ながらこちらの場合も前者の実験群と後者の対照群では語彙習得量に違いがなく，カバー率が良ければ語彙習得も良くなることが分かった。日本語字幕と英字幕の連続的な使用の効果がなかった理由は，1つ目の研究設問同様，字幕の有無に関わらず，言語面に向けられた注意量が同じであることや，使用した動画が音声的に難しくなかったことが考えられる。さらに1つ目の研究設問のグループの記述統計量も踏まえて，3つ目の研究設問グループ間で偶発的語彙習得の量に違いがなかったもう一つの理由として，視聴回数と回数に応じた集中度合いの要因も考えられた。今回研究設問の対象となった4グループ（Exp1, Exp2, Con1, Con2）はすべて動画を2回視聴している。1回目の視聴時点では，日本語であれ英語であれ字幕のあるグループのほうが内容理解度が有意に良かったので，偶発的語彙習得の量もまた多かったかもしれない。しかし，1回目に字幕があったグループの参加者は，最初の視聴で既に十分な理解を得られているため，2回目の視聴時は動画視聴にあまり集中せず，逆に1回目に字幕のなかったグループの参加者は内容理解度が低く，2回目の視聴も集中して取り組んだ可能性が考えられた。そのため，結果的にすべてのグループで似たような語彙習得量になったのかもしれない。繰り返し学習することが好影響を与えるのは語彙習得も例外ではなく，語彙に何回も触れるほど習得の可能性も高まる（Uchihara et al., 2019）。そのため，クラス内や自主学習の環境では第二言語で動画を観るときには，それが本実験で使用したような短いマテリアルである場合，繰り返し視聴することが良いだろう。その際，英字幕や日本語字幕をうまく活用していくことに語彙習得にとどまらない意義があると考える。ただし，視聴回数が増えるにしたがって，学習者の集中力やモチベーションが低下しないように，動画を立て続けに視聴するというよりは，視聴回数に応じて注目するべきポイントを変えたり，言語活動を間に含めたりするなどの工夫が必要だと考える。もちろん，字幕の効果的な使用法に関する研究は今後更に必要であるのは言うまでもない。今回，統計モデルに含むことのできなかった字幕の有無とカバー率の交互作用の問題は興味深い。つまり，動画内に出てくる単語の95%以上を知っている場合に更に英字幕があることで偶発的語彙学習が英字幕のない場合より良くなるのか，研究の価値がある。更に，学習者の視聴回数に応じたふるまいの違いについては視線計測機などの使用が見込まれる（Teng, 2021）。ただし，ヘッドマウント型の視線計測機は参加者になるべく顔を動かさないことを求める厳密な実験となるため，参加者の現実世界における振る舞いまでは考慮に入れることができないかもしれない。参加者へのインタビューやアンケートを通じて，動画視聴に関する率直な意見を聞くのもいいだろう。この点において，実際の教育現場の先生方によるアクションリサーチ（Burns, 2011）にも期待したい。

教室内外で使用できる技術は字幕に限らず，学習に有効とされる様々なコンピューター支援言語学習の道具がある（Chun, 2011）。第二言語視聴と字幕に関する本研究の結果が，現場の指導者の方々の指導レパートリーの一つになったり，研究者が更なる研究を考案したりすることに寄与できることを願う。

謝辞

本研究の機会をくださった公益財団法人 日本英語検定協会のご担当者様，選考委員会の先生方に，厚くお礼を申し上げたいと思います。中でも，研究助言者の西垣知佳子先生（千葉大学），私が修士課程に在

籍していたときの指導教官であった根岸雅史先生(東京外国語大学),そして私の師匠である内原卓海先生(東北大学)の3人の先生方には非常に貴重で有益なご助言,ご指導を賜りましたこと,感謝のしようがありません。また,本実験の参加者を集めるのに協力してくださった各先生方にもこの場をお借りしてお礼申し上げます。実験に参加してくださった参加者の皆様にも多大なお礼を忘れません。

引用文献

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior research methods, 52*(1), 388-407. https://doi.org/10.3758/s13428-019-01237-x

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLoS ONE, 11*(6). Article e0158409. https://doi.org/10.1371/journal.pone.0158409

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed). John Wiley & Sons.

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science, 4*(1), 1-19. https://doi.org/10.1177/2515245920960351

Burns, A. (2011). Action research in the field of second language teaching and learning. In E. Hinkel (Ed), *Handbook of research in second language teaching and learning volume 2.* Routledge.

Charles, T., & Trenkic, D. (2015). Speech segmentation in a second language: The role of bimodal input. In Y. Gambier, A. Caimi & C. Mariotti (Eds), *Subtitles and language learning: Principles, strategies and practical experiences* (pp. 173-198). Peter Lang.

Clenton, J., & Booth, P. (2021). *Vocabulary and the four skills: Pedagogy, practice, and implications for teaching vocabulary.* Routledge.

Chacón-Beltán, R., Abello-Contesse, C., & del Mar Torreblanca-López, M. (2010). Insights into non-native vocabulary teaching and learning. Multilingual Matters.

Chun, D. M. (2011). Computer-assisted language learning. In E. Hinkel (Ed), Handbook of research in second language teaching and learning volume 2. Routledge.

Cobb, T. (n.d.). Compleat Web VP (Version 2.6) [Computer Program]. https://www.lextutor.ca/vp/comp/

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*(3), 369–382. https://doi.org/10.1177/0267658312443651

Dizon, G., & Thanyawatpokin, B. (2021). Language learning with Netflix: Exploring the effects of dual subtitles on vocabulary learning and listening comprehension. *Computer Assisted Language Learning Electronic Journal, 22*(3), 52-65.

Durbahn, M., Rodgers, M., & Peters, E. (2020). The relationship between vocabulary and viewing comprehension. *System, 88,* 1-13. https://doi.org/10.1016/j.system.2019.102166

Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-26). Multilingual Matters.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching.* Multilingual Matters.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed). Sage, Thousand Oaks CA.

Frumuselu, A. D., De Maeyer, S., Donche, V., & Plana, M. D. M. G. C. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education, 32,* 107-117. https://doi.org/10.1016/j.linged.2015.10.001

Gamer, M., & Lemon, J. (2019). irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1, <https://CRAN.R-project.org/package=irr>.

Garcia, G. D. (2021). *Data visualization and analysis in second language research.* Routledge.

Graham, S., & Zhang, P. (2021). Vocabulary and listening: Current research, tools, and practices. In J. Clenton & P. Booth (Eds), *Vocabulary and the four skills: Pedagogy, practice, and implications for teaching vocabulary.* Routledge.

Hao, T., Sheng, H., Ardasheva, Y., & Wang, Z. (2022). Effects of dual subtitles on Chinese students' English listening comprehension and vocabulary learning. *The Asia-Pacific Education Researcher, 31*(5), 529-540. https://doi.org/10.1007/s40299-021-00601-w

Hinkel, E. (2011). *Handbook of research in second language teaching and learning volume 2.* Routledge.

Hsieh, Y. (2020). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Computer Assisted Language Learning, 33*(5-6), 567-589. https://doi.org/10.1080/09588221.2019.1577898

Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In C. P. Robinson (Ed), *Cognition and second language instruction* (pp. 258-276). Cambridge University Press.

Jelani, N. A. M., & Boers, F. (2018). Examining incidental vocabulary acquisition from captioned video: Does test modality matter? *ITL-International Journal of Applied Linguistics, 169*(1), 169-190. https://doi.org/10.1075/itl.00011.jel

引用文献

Kang, E. (2019). Enhancing Korean EFL learners' vocabulary learning and listening comprehension through video captions. *STEM Journal, 20*(2), 91-108. doi: 10.16875/stem.2019.20.2.91

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge University Press.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*(3), 399-436. https://doi.org/10.1111/j.0023-8333.2004.00260.x

Mayer, R. E. (2005). Introduction to multimedia learning. In R. E. Mayer (Ed), *The Cambridge handbook of multimedia learning* (pp. 1-18). Cambridge University Press.

Mayer, R. E. (2021). *Multimedia learning* (3rd ed). Cambridge University Press.

Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching, 55*(2), 163-192. doi:10.1017/S0261444821000501

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. Language *Learning & Technology, 18*(1), 118-141. http://llt.msu.edu/issues/february2014/monteroperezetal.pdf

Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning, 31*(1-2), 1-26. https://doi.org/10.1080/09588221.2017.1375960

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720-739. http://dx.doi.org/10.1016/j.system.2013.07.013

Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL quarterly, 53*(4), 1008-1032. https://doi.org/10.1002/tesq.531

Plonsky, L. (2015). *Advancing quantitative methods in second language research.* Routledge.

Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal, 100*(2), 538-553. https://doi.org/10.1111/modl.12335

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912. https://doi.org/10.1111/lang.12079

Plonsky, L., & Oswald, F. L. (2015). Meta-analyzing second language research. In L. Plonsky (Ed). *Advancing quantitative methods in second language research* (pp. 106-128). Routledge.

R Core Team (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Raine, P. (2013, January). Incidental increase in depth of vocabulary knowledge through the viewing of subtitled, authentic videos [Proceedings]. JALT 2012 Conference. https://www.researchgate.net/publication/315658011_Incidental_Increase_in_Depth_of_Vocabulary_Knowledge_Through_the_Viewing_of_Subtitled_Authentic_Videos

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research_. Northwestern University, Evanston, Illinois. R package version 2.3.6, <https://CRAN.R-project.org/package=psych>.

Robinson, P. (2001). *Cognition and second language instruction.* Cambridge University Press.

Rodgers, M. P. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions* [Doctoral dissertation, Victoria University of Wellington]. http://researcharchive.vuw.ac.nz/handle/10063/2870

Rost, M. (2016). *Teaching and researching listening* (3rd Ed). Routledge.

Schmitt, N. (2010). Key issues in teaching and learning vocabulary. In R. Chacón-Beltán, C. Abello-Contesse, M. del Mar Torreblanca-López (Eds), *Insights into non-native vocabulary teaching and learning.* Multilingual Matters.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913-951. https://doi.org/10.1111/lang.12077

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26-43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Ted-Ed. (2016, April). The evolution of the book [Video]. YouTube. https://www.youtube.com/watch?v=_YqYtdPUis4

Teng, F. (2018). Incidental vocabulary learning for primary school students: the effects of L2 caption type and word exposure frequency. *The Australian Educational Researcher, 46,* 113–136. https://doi.org/10.1007/s13384-018-0279-6

Teng, M. F. (2021). *Language learning through captioned videos: Incidental vocabulary acquisition.* Routledge.

Teng, M. F. (2022). Incidental L2 vocabulary learning from viewing captioned videos: Effects of learner-related factors. *System, 105,* 1-12. https://doi.org/10.1016/j.system.2022.102736

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning, 69*(3), 559-599. https://doi.org/10.1111/lang.12343

Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning.* Springer.

Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L2 and L2 listening comprehension: The same or different from reading comprehension? *Applied linguistics, 34*(4), 457-479. https://doi.org/10.1093/applin/ams074

引用文献

Wang, A., & Pellicer - Sánchez, A. (2022). Incidental vocabulary learning from bilingual subtitled viewing: An eye - tracking study. *Language Learning, 72*(3), 765-805. https://doi.org/10.1111/lang.12495

Webb, S. (2007). The effects of repetition on vocabulary knowledge. Applied Linguistics, 28(1), 46-65. https://doi.org/10.1093/applin/aml048

Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed), *The Routledge handbook of vocabulary studies* (pp. 225-239). Routledge.

Webb, S. (2020). *The Routledge handbook of vocabulary studies.* Routledge.

Webb, S., & Nation, P. (2017). *How vocabulary is learned.* Oxford University Press.

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

Wei, R., & Fan, L. (2022). On-screen texts in audiovisual input for L2 vocabulary learning: A review. *Frontiers in Psychology, 13.* doi: 10.3389/fpsyg.2022.904523

Wesche, M. & Paribakht, T. (2010). *Lexical Inferencing in a First and Second Language: Cross-linguistic Dimensions.* Multilingual Matters. https://doi.org/10.21832/9781847692245

Wi, I., & Boers, F. (2023). Sequential Use of L1 and L2 Captions: Exploring the Benefits for Vocabulary Acquisition. *TESOL Quarterly.* https://doi.org/10.1002/tesq.3243

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A Grammar of Data Manipulation_. R package version 1.1.2, <https://CRAN.R-project.org/package=dplyr>.

A 研究部門 ● 報告Ⅳ ● 英語能力テストに関する研究

複数テキスト読解でのつながり形成が日本人EFL学習者の統合的理解に与える影響の検証

研究者:三上　洋介　埼玉県／駿河台大学　助教
（申請時:東京都／東京都立国分寺高等学校　教諭・茨城県／筑波大学大学院　在籍）
《研究助言者:小泉 利恵》

概要

本研究の目的は，複数テキストを読解した際のテキスト間のつながり形成と読後の統合的理解の関係を，即時処理と事後処理の2つの測定方法で調べることである。実験では，4つの異なる関係性で構成される複数テキストセットが用意され，協力者はテキストセットを読みながら，テキスト間のつながりを特定し記述することが求められた（即時処理）。読解後，統合的な理解を測定するためライティングによる要約設問が与えられた（事後処理）。その結果，総つながり数に関しては，テキスト要因での影響は見られなかった。一方，熟達度の主効果が有意で熟達度が高い人はつながり形成数が多く，熟達度要因はつながり形成に影響していた。また個別に形成されたつながりの数と要約ライティングスコアにおいて，正の相関関係が見られた。つまり複数テキストを読解した際のテキスト間のつながり形成と読後の統合的理解には関連性が見られた。

1 序論

近年，単一テキストだけでなく，複数のテキストを読み比べて，特定のトピックを理解したり学習したりすることは日常的になっている（Stang Lund et al., 2019）。第二言語（L2）学習者にとっても例外ではない。対象言語のインプット量が限定的である日本において，インターネット等を活用して複数の英語テキストを読解することは，物理的制約を乗り越え言語インプット不足を解消するための有効な対策である（日本学術会議, 2016）。

しかし複数テキストの理解は単一テキスト理解よりも複雑である（Perfetti et al., 1999）。複数テキストを理解するためには，個々のテキストを理解するだけでなく，それらのつながりを形成する必要があるためである（List et al., 2021a）。例えば個々のテキストは独立しており，それぞれの関係性は明示的に述べられておらず，読み手自身によって推論し，個々のテキストを整合性のあるものに統合して解釈しなければならないため，第一言語（L1）話者だけでなく，L2学習者にとってもつながり形成は難しい（Schedl et al., 2021; Wineburg, 1991）。

外国語としての英語（English as a foreign language; EFL）を学ぶ日本人にとって，複数テキストの理解はなおさら困難であることが想定されるが，複数テキスト処理メカニズムの検証は限定的である。そこで本研究はこのテキスト間のつながり形成処理と統合的理解の関係に焦点を当てる。本研究の目的は，日本人英語学習者が複数テキストを読解する時に形成されるテキスト間のつながりと，読後の統合的理解の関係を，テキスト間の関係性が異なる構造の複数テキストを用いて2つの測定方法（即時処理と事後

処理）で明らかにすることである。

2 先行研究

2.1 主体的・対話的で深い学びを促すための読解活動

近年，新しい時代を生き抜く力を養うため，教育において主体的・対話的で深い学びが推奨されている。主体的な学びとは，「学ぶことに興味や関心を持ち，自己のキャリア形成の方向性と関連付けながら，見通しを持って粘り強く取り組み，自己の学習活動を振り返って次につなげる」ことである（文部科学省，2018，p.128）。また対話的な学びとは，「子供同士の協働，教職員や地域の人との対話，先哲の考え方を手掛かりに考えること等を通じ，自己の考えを広げ深める」ことである（文部科学省，2018，p.118）。この主体的で対話的な深い学びを生じさせる上で参考になるのが，ブルームの教育目標の分類学（Bloom taxonomy）である。この分類では，学習者は学習題材を分析・評価することにより高度な学びが生じると定義している（Bloom et al., 2001）。特に「分析」は，情報を分解し，共通点や相違点を見つけてつながりを形成すること，「評価」は，立場や決定を正当化すること，と定義されている。

こうした読み手の分析や評価を促し，学びを誘発する手段の一つに，複数テキストを読み比べる活動がある。例えば，内容のくいちがうテキストを読み合わせるとき，読み手はそれぞれの内容を額面通り受け取ることが出来ない（Rouet et al., 2021）。このくいちがいを解消し，整合性のある理解を得るために，読み手は内容を分析することや，著者や出版者情報などのソース情報を評価することへの認知的努力が求められる（Braasch & Bråten, 2017）。Rouet et al.（2021）によると，対立する２つのテキストを読み手に与え，情報源であるソース情報の信頼性を評価させるタスクを与えたとき，ソース情報の統合的な理解が促進された。つまり複数のテキストを読み比べさせることにより，読み手にテキスト内容やソース情報を評価分析することを促すことが可能である。結果として，内容が対立する複数テキストを読解することによって，読み手は柔軟かつ多面的な複数の視点から対象を理解することが可能になる（Bråten & Braasch, 2018）。

また複数テキストを理解することは，脈絡のない個々のテキストを，学習者が自らの意思で関連付け，整合性のある統合された理解を構築することである（Primor et al., 2021）。この時，複数テキストの読み手は，受け身ではなく主体的に意味を解釈し，新しい理解を作り出す学習が求められる。加えて，あるテーマについて対立するテキストを読解するとき，内容やソース情報を判断する必要が生まれるため，単に単一のテキストを読むよりも，読み手の中に主体的で対話的な学びが生じやすい。このように，複数テキストを読み，統合的な理解を構築する行動は，主体的・対話的で深い学びの促進に有用と言える。

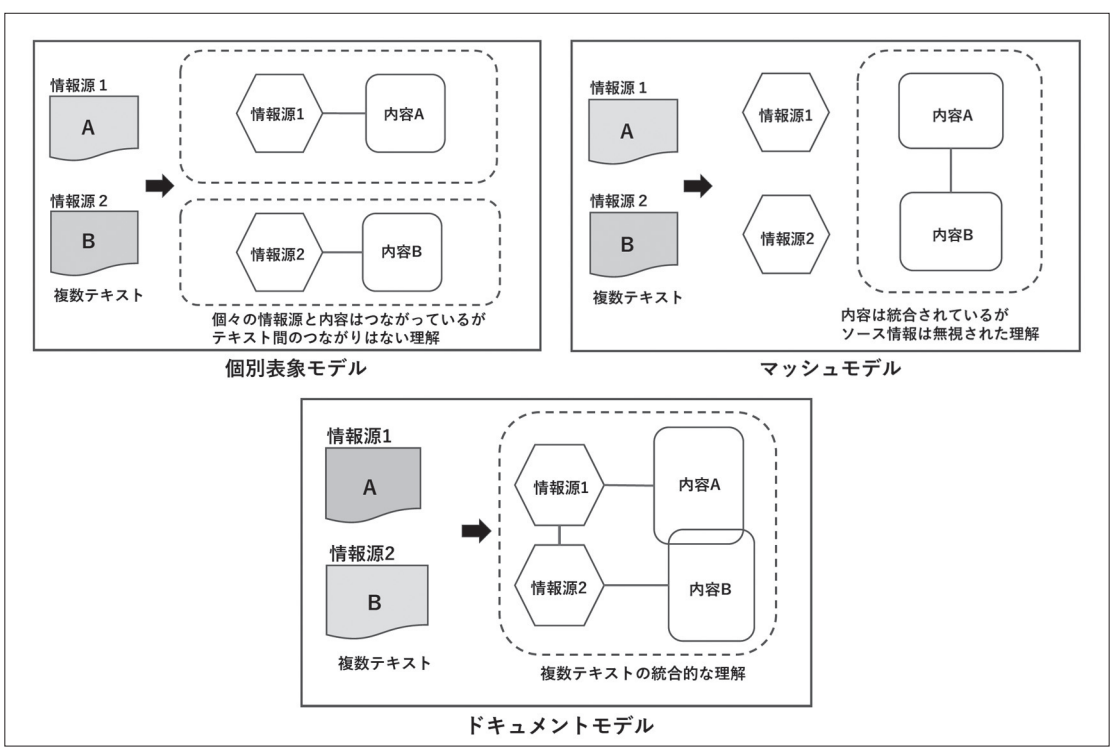
2.2 読解理論（DMF, D-ISC）と統合的理解

複数テキストを統合して理解し，心内に整合性のある表象を作ることを，統合的心的表象（integrated mental representation）の構築，または統合的理解（integrated understanding）の構築という（List et al., 2021a）。つまり複数テキストを理解するということは，あるテーマに対し，個々のテキストを越えて一貫した新しい解釈を作り出すことである。

複数テキスト読解処理を説明するのがドキュメントモデルフレームワーク（Document Model Framework; DMF）である。 このモデルでは，読み手は内容と，情報源であるソース情報を心の中に表象すると仮定する。そして複数テキストの理解は，この内容の統合，ソースの表象，内容とソースのつながりの3要素で構成され，それぞれの要素の結びつきにより，読み手は異なった複数テキスト理解を形成する（Perfetti et al., 1999）。 内容統合とソース，内容のつながりの程度により，複数テキスト理解の状態はそれぞれ（1）個別表象モデル，（2）マッシュモデル，（3）ドキュメントモデル（DM）と呼ばれる（Britt et al., 1999; 図1）。

まず「個別表象モデル」では，読み手は個々のテキスト同士をつなげずに独立した表象として理解を構築する。つまり複数テキストの読み手が，個々のテキストの内容を1つの心的表象に統合しない場合，テキストごとの心的表象が別々に形成される。この時，個々のテキストは関連付けられず，個別に理解されている。この状況は，最初のテキストからの学習が，次のテキスト読解に活かされないため好ましくない。一方，「マッシュモデル」は，個別表象モデルと逆の性質を持つ。テキストから学習したすべての情報は統合されるが，ソース情報（ソースの信頼性など）は無視される。個々のソース情報を考慮せず，内容が統合されるため，読み手の記憶の中でソース情報を利用できない状態である。テキスト間の情報の重複度が高い場合や，間を置かずに複数のテキストを読む場合に，このマッシュモデルは生じやすい。状況理解が促進される利点がある一方，情報の真偽や評価の判断が難しくなる欠点がある。このように個別表象モデルは内容統合がされず，マッシュモデルはソース情報の統合が欠落しており，いずれも複数テキストの統合的理解には不十分である（List & Alexander, 2019）。

最も望ましい状態はDMである。DM は個々のテキスト内容を1つの心的表象に統合しつつ，ソース情報とのつながりも形成される，統合的に理解された状態である。DMにおいては，個別表象またはマッシュモデルの一方の側を選択せず，内容のくいちがう情報を表象できる（Schoor et al., 2023）。つまり複数テキストを読解する目的は，個々のテキストを理解し，それぞれをつなげて複数テキストを統合的に理解し，整合性のある心的表象である DM を形成することである。



■図1：複数テキスト理解の状態（Britt et al., 1999; Saux et al., 2021に基づく）

2.3 対立テキストと統合的理解

複数テキストは，テキストの関係性（text relations）により，大きく二つに分類される。一つは同じテーマを異なる視点で述べて，補完関係にある「相補的テキスト」であり，もう一つは内容がくいちがう「対立的テキスト」である（e.g., Barzilai et al., 2018; Perfetti et al., 1999）。Ferguson et al.（2013）によると，複数テキストの統合的な心的表象である DM の形成には，特に対立的テキストが役立つ。

D-ISC（discrepancy-induced source comprehension）モデルはそのメカニズムを説明する。D-ISC モデルによると，相補的テキストと対立的テキストでは読み手の読解処理は異なる（Braasch & Kessler, 2021）。まず相補的テキストの場合，同じトピックについて異なる視点で書かれており，意味の重複度が高いため，読み手は自動化された状態でテキスト情報を心的表象上で更新する（Kintsch, 1998）。一方，対立的テキストの場合，読み手は情報のくいちがいに出会うと，認知的不均衡（cognitive disequilibrium）を経験する。対立関係を理解し一貫した心的表象を構築しようと，読み手はより努力を要する方略的処理に移行し，ソース情報に注意を向けて内容理解が深まると仮定する（Braasch & Bråten, 2017; van den Broek & Kendeou, 2015）。

この対立的テキストのくいちがいは，内容の整合性がある心的表象である DM の形成を促す。例えば，対立情報を読む時，読み手はよりソース情報に読み戻り，長時間凝視する。また対立情報を読んだ場合，要約ライティングにソース情報を含める。そして対立情報を読んだ方が，ソース情報の記憶が強化される（Braasch et al., 2012）。つまり対立情報を読むと，読み手は内容のくいちがいを解消しようとしてソース情報への注意を向ける。その結果，テキスト間のソース情報のつながり形成が促進され，整合性ある DM の形成に役立つ。このように複数テキストを統合的に理解するためには，個々のテキスト理解に加え，テキストの関係性を整理しつつ，テキスト間の内容とソース情報のつながりを形成する必要がある（Britt et al., 1999）。

しかしStadtler et al.（2020）によると，複数テキストの理解は，L1話者にとっても困難な場合があり，往々にして読み手は単一の整合性のある心的表象を形成することが難しい。その理由の一つが，テキスト間のつながり形成の難しさである。単一テキストはテキスト内の情報のつながりが明示的に説明されているが，複数テキストは物理的に隔離され，個々のテキスト同士についてのつながりは明示的に書かれておらず整合性がない状態で存在している。このため複数テキストの読み手は，整合性がない状態で存在している個々のテキストを，努力してつなげ，整合性のある理解を形成しなければならない（Primor et al., 2021; Rouet et al., 2016）。

例えば対立する歴史的事実についての複数テキストにおいて読み手に生じる最も容易なことは，テキスト間のつながりを形成せず，個々のテキストを「島々」のようにそのままにすることである（Wineburg, 1998）。よって複数テキストを読み，統合して理解し整合性のある意味を構築することは実は難しく，読み手にとって大きな認知的達成である。

2.4 つながり形成

2.4.1 つながり形成と統合レベル

List（2020）によると，複数テキストにおける統合は3種類ある：（1）テキスト上の情報と読み手の既存知識との統合，（2）単一テキスト上に提示されている情報同士の統合，そして（3）一見無関係に見える複数テキスト間の統合に区別される。この三番目の統合はテキスト間統合（intertextual integration）と呼ばれ，読み手にとって最も難しい統合である（List, 2020）。本研究の統合は，複数テキスト理解を成功するために重要とされる「テキスト間統合」を指す。

このテキスト間統合は，複数テキストにのみ見られる読解処理であり，テキスト間情報のつながり形成や結合処理を含み，意味構築に重要といわれ，複数テキスト理解の核である（Primor et al., 2021）。このように，複数テキストの統合的な心的表象を形成し統合的理解を成功させるためには，個々のテキスト理解だけでは十分ではなく，テキスト間のつながりを特定し，統合する処理が必要である（Bråten et al., 2011; Britt et al., 2013）。

2.4.2 つながりの種類

List et al.（2021a）によると，テキスト間統合のうち，複数テキスト間の類似性や差異性を認識し，つながりを形成し，テキスト間の関係性を精緻化する特殊な読解行動を「つながり形成」と呼ぶ。このつなが

り形成には，異なる種類とレベルがある。まず，複数の説明文を読む時，読み手は少なくとも3つの特定性を持つつながりをテキスト間で形成する。List and Alexander（2019）は，証拠的つながり，テーマ的つながり，そして文脈的つながりを定義した（表1）。証拠的つながり，テーマ的つながりは形成されやすい一方，メタ的認知を必要とする文脈的つながりはもっとも形成されにくい（List et al., 2021a）。

■表1: 複数テキスト間のつながり形成の特定性（List et al., 2019を一部修正）

特定性	定義	例
証拠的つながり （evidentiary connection）	根拠や情報の入手方法など テキスト間の特定情報への働きかけ	“両方の記事では学校建設の利点として， 教育の質を高めることが共通している。”（根拠） “調査は市長室によるアンケートと地元紙の インタビューでデータは収集されたのか。”（方法）
テーマ的つながり （thematic connection）	テキスト間のメインアイディアを 比較したもの	“両方の記事では住民はショッピングモール建設 に賛成している。”
文脈的つながり （contextual connection）	著者の信頼性や専門的視点といった メタテキスト特徴のつながり	“Local Voice Timesは地元誌だから 地元住民側の主張を載せているが， Business Daily Pressは経済界の意向が 反映されているのかな。”（下線部がメタテキスト特徴）

2.4.3 統合レベル

List et al.（2021b）によると，テキスト間のつながりには統合の深さに応じてレベル1から4で構成される，進行性の統合レベルが仮定されている（表2）。レベル1は単につながりに気づいている低次の統合レベルであり，レベル4は推論生成を含む高次のつながりである。レベル3の同時的関係が，DMFのテキスト間統合にもっとも近く，レベル4の関係的精緻化はL1学習者でも最も形成が困難である。つまり，つながり形成には単に情報を選択しつなげる下位の処理と，より精緻化された上位の処理がある。

■表2: 統合の4段階（List & Alexander, 2019を一部修正）

Level	統合	定義	例
1	関連的特定 （relational identification）	テキスト間の潜在的重複やつながり への気づきや注意	“2つの記事は施設建設に対する住民の意見に ついてかな。”
2	個別表象 （separate representation）	個々のテキスト内で提示された情報の 潜在的な関係の個別の理解	“1つ目の記事では学校建設によってクラスサイズ が小さくなり，そして2つ目の記事では同時に野球 やサッカーの練習ができるようになると言っている。”
3	同時的関係 （simultaneous relation）	接続詞を用いた単一的，組み合わせの 陳述によりテキスト同士をつなげること	“学校建設に賛成する根拠の観点は，学習面と 運動面への好影響だな。”
4	関係的精緻化 （relational elaboration）	前のレベルで特定された関係を 精緻化しより説明すること	“新しい学校建設の利点は，要は教室の密度が 減って容積が増えるのだな。”

2.5 テキストの関係性とつながり形成

Braasch et al.(2016)はD-ISC モデルを土台とするトレードオフ仮説を提唱し，テキストの関係性はつながり形成に影響すると仮定した。このトレードオフ仮説によると，テキストの関係性と形成されるつながりの種類はトレードオフの関係がある。相補的テキストは主張やサポート情報において意味の重複が多く，読み手は自動的に情報を処理しやすいため，証拠的・テーマ的つながりの形成が促進される一方で，メタ的認知が求められる文脈的つながりの形成は阻害される。逆に対立的テキストの場合，読み手は内容を額面通り受け取れないため，よりソース情報に注意を払うことで文脈的つながりを一層形成させるが，証拠的・テーマ的つながりへの注意が制限され，その形成は抑制される(Braasch et al., 2016)。L1の実証研究では，このトレードオフ仮説が支持された(List et al., 2021a)。

一方でトレードオフ仮説が当てはまらない可能性もある。つながりを形成する処理では，特に対立的テキストで認知的負荷が高いため，L1の読み手は対立的テキスト間のつながりを無視する。このためくいちがいの認知的不均衡を経験せず，心的表象への統合を控えることがある(Stadtler et al., 2020)。またこれまで検証はされていないが，対立的テキストでも証拠的・テーマ的つながりが促進される可能性もある。このように L1 研究においてテキストの関係性が複数テキストのつながり形成に与える影響については，一致した見解に達していない。

Braasch and Kessler（2021）によると，L1に比べて，L2学習者は対立的な複数テキストを読むとき，情報のくいちがいを認知することが苦手で，くいちがいを無視し，テキスト間をつなげる処理を行わない傾向がある。EFL学習者にとってL1研究結果があてはまらない可能性があり，特に熟達度の低いEFL学習者は，対立的テキストのつながり形成に苦労する可能性がある。またEFL学習者の複数テキスト理解は，DMまで到達せず，個別表象モデルやマッシュモデルにとどまる可能性が高いかもしれない。

2.6 L2 学習者の熟達度が与える影響

複数テキストの読解処理は，単一テキストの読解処理に比べて複雑である(Perfetti et al., 1999)。さらに上述のように L1話者にとっても複数テキストの読解中のつながり形成は困難であることを踏まえると，L2学習者の熟達度が低い場合，より複数テキストの読解中のつながり形成に困難を抱える可能性がある。

それは読解処理の点から説明できる。まず読解処理は単語認識や統語処理などの下位処理と，推論生成などの上位処理で構成されるが，L2学習者は下位処理に認知資源を割かれるため推論生成などの上位処理が少ない(Horiba, 2013)。つまり認知容量に余裕のある高熟達度学習者は上位処理に従事できる一方，低熟達度学習者は下位処理にとどまる。例えばL2の複数テキストを読解するときの読解方略研究では，高熟達学習者と低熟達学習者が対立的テキストを読解した場合，高熟達度学習者のほうが複数テキストの内容全体を把握しようと務めるのに対し，低熟達度学習者は単語レベルの理解に固執していた(Plakans, 2009)。このように高熟達度学習者のほうが複数テキストを俯瞰的にとらえようとするため，テキスト間のつながりを特定し，形成しやすい可能性がある。

一方，複数テキストの読解中につながりを特定させるタスクを与えた場合には，L2の熟達度の高低は複数テキストの読解処理に影響しないという研究もある。つながりの数を直接調べた研究ではないが，高熟達度群と低熟達度群に対し，複数テキスト読解中につながりを特定させるタスクを与え，その後，List et al.(2021a)と同じ採点基準を用いた統合レベルを測る要約ライティングと，統合理解問題を課したところ，両方のスコアに熟達度の差はみられなかった(三上, 2022)。このようにL2学習者の熟達度が，複数テキスト読解中のつながり形成に影響するという研究と，影響しないという研究があるが，すべて事後処理で測られたもので，これまで即時処理でのつながり形成を直接検証した例は見られない。要約ライティングなどの事後処理を検証するだけでは，L2学習者の読解処理を十分に把握できないため，事後処理に加えて即時処理も対象とすることで，学習者の複数テキストに対する読解処理を相互補完的に把握することが可能になる。

2.7 つながり形成と統合的理解の関係

L1研究では，複数テキストの読解中に形成されたつながり数と，統合的な理解に相関関係が報告されている。List et al.（2021a）では，読み手に2つのテキストのつながりを特定させた後，テキストを参照できない状態で統合的な要約ライティングを課した。その結果，読解中に形成された総つながり数と，要約ライティングのスコアには弱い相関関係（r=.22）がみられた。そして即時処理でのつながり形成を多く行う読み手は，事後処理での要約ライティングスコアが高い傾向が若干あった。つまり即時処理でのつながり形成と，事後処理での要約ライティングスコアには弱い相関関係があった。

L2研究では，即時処理ではないが，検証例はある。複数テキストの読解中，つながり形成を促すテキストハイライティングタスク介入を経験した協力者は，介入が行われなかった協力者よりも要約ライティングのスコアが高かった（三上, 2022）。この結果から，EFL 学習者においても，即時処理でのつながり形成と事後処理での複数テキストの統合的理解には関係性があると想定される。

しかし，L2学習者の特性から，複数テキストにおけるつながり形成と統合的理解の関係は L1と異なる結果も想定される。例えば，L1研究とL2研究では，より認知的負荷の高い対立的テキストの処理で異なる結果が出ている。L1研究で，List et al.（2021a）は2つの相補的または対立的テキストを読んだ後，統合的な心的表象を反映させるため，テキストを参照させずに要約ライティングを課した。その結果，対立的テキストよりも相補的テキストのほうが要約スコアは高かった。また読解中にテキスト間で形成されるつながりも，対立的テキストよりも相補的テキストのほうが多かった。つまり即時処理と事後処理の測定方法で，L1話者は対立的テキストの対処に苦労していた。

一方，L2研究では対立的テキストをよりうまく対処した研究がある。Qin and Liu（2021）によると，L2学習者に対し，2つの長い対立的または相補的テキストを読解後，テキストの参照を可能とした要約ライティングを課した場合，対立的テキストのほうが要約スコアは高かった。またテキスト間のつながりは，対立的テキストのほうが，特に証拠的つながりであるデータ情報が多かった。つまりL2学習者にとっては，対立的テキストのほうが，相補的テキストよりも対処しやすかった。

これは前述のL1研究と異なる結果である。しかも，L１研究（List et al., 2021a）ではテキスト参照は不可であり，L2研究（Qin & Liu, 2021）ではテキスト参照は可とされていたので，ライティングタスク指示の違いが結果に影響した可能性もある。このようにQin and Liu（2021）は要約ライティング時にテキスト参照を可能としたが，要約ライティングで統合的な心的表象を測るためには，テキスト参照を不可とさせる必要がある（Britt et al., 1999; Perfetti et al., 1999）。よって読み手の統合的理解を測定するために，この実験手法は適切ではなかった可能性がある。加えてQin and Liu（2021）は事後処理のみを測定対象としているため，L1研究（List et al., 2021a）との厳密な比較は難しい。

このように，L2学習者にとっては，相補的や対立的といったテキストの関係性によって，つながりの数と統合的理解の関係はL1と異なる結果になる可能性がある。しかしL1とL2の先行研究では，実験手法やマテリアルが異なっているため，研究結果を比較検証することは難しい。もしつながり形成と統合的理解の関係性を先行研究に近い条件で検証できれば，L2学習者が複数テキスト理解の際に何につまずくのかが明らかにできるだろう。

2.8 本研究

先行研究の限界点は，次の3点である。

1. L1話者に比べて，L2学習者は対立的な複数テキストを読むとき，情報のくいちがいを認知することが難しく，くいちがいを無視し，テキスト間をつなげる処理を行わない傾向がある（Braasch & Kessler, 2021）。このためEFL学習者にとってL1研究の結果があてはまらない可能性があり，熟達度の低いEFL学習者は，対立的テキストのつながり形成に苦労する可能性があるが，これまで検証されていない（2.4参照）。

2. L2学習者の熟達度が，複数テキストを読解するときのつながり形成に影響するという研究と，影響しないという研究があり，さらなる検証が必要である。さらに，これまでは事後処理で検証され，即時処理でのつながり形成を直接検証した例は見られない（2.5参照）。
3. L1研究では，事後処理と即時処理の両方で，複数テキストの読解中に形成されたつながり数と，要約ライティングスコアによる統合的な理解に相関関係が報告されているが，L2研究では，事後処理と即時処理の両方に焦点を当てた検証例はない（2.6参照）。

上述の限界点を受けて，本研究は対立的・相補的な情報を含む複数テキストに焦点を当てる。本研究の目的は，複数テキストを読解した際のテキスト間のつながり形成と読後の統合的理解の関係を，即時処理と事後処理の2つの測定方法で調べることである。2つの観点で調べることにより，統合的理解に成功しているEFL学習者は読解中にどのような読解処理を行っているかを明らかにできる。

本研究の新規性は次の2点である。

1, EFL研究において，テキスト間の関係性を統制し，読解中のつながり形成と，読後の統合的理解を検証した最初の研究である。
2, 熟達度の異なるEFL学習者が複数テキストを読解中，何につまずいているのかをつながり形成の視点から調べた最初の研究である。

先行研究から次の予測が導き出される（表3）。

■表3: 本研究の予測

	予測1	予測2	予測3
分析1-1 総つながり数と テキスト関係	L1研究と同じ結果を想定する場合，総つながり数はテキストの関係性によって影響を受ける。特につながり形成数が相補的テキストよりも少ない 。	総つながり数はテキストの関係性によって影響を受けるが，EFL学習者は対立的テキストを比較的うまく処理するため，つながり形成数が相補的テキストよりも多い 。	総つながり数はテキストの関係性によって影響を受けない。
分析1-2 総つながり数と 熟達度	L2学習者の特性を踏まえると，熟達度はつながり形成に影響する。	熟達度はつながり形成に影響しない。	
分析2 タイプ別のつながり数と テキスト関係	L1研究と同じ結果を想定する場合，テキストの関係性によって，形成されるつながりの特定性は影響を受ける（トレードオフ仮説を支持）。	L2学習者特有の要因を想定すると，テキストの関係性によって，形成されるつながりの特定性は影響を受けない（トレードオフ仮説を不支持）。	
分析3 テキスト関係と 要約ライティングの関係	テキスト関係は要約ライティングスコアに影響する。	テキスト関係は要約ライティングスコアに影響しない。	
分析4 テキスト間つながりと 統合的理解の関係	個別に形成されたつながりと要約ライティングスコアは，L1と同様に相関関係がある。	つながりと統合的理解には相関関係は見られない。	

検証課題（RQs）は以下の通りである。

RQ1 テキスト間の関係性を統制した複数の説明文テキストを読解するとき，形成されるテキスト間のつながり総数，異なる特定性のつながり数はテキスト要因において違いがみられるか（分析1, 分析2）。

RQ2 テキスト間の関係性を統制した複数の説明文テキストを読解するとき，読解中に形成されるテキスト間のつながりは読後の要約ライティングにおける統合的理解と関係があるか（分析3, 分析4）。

3 方法

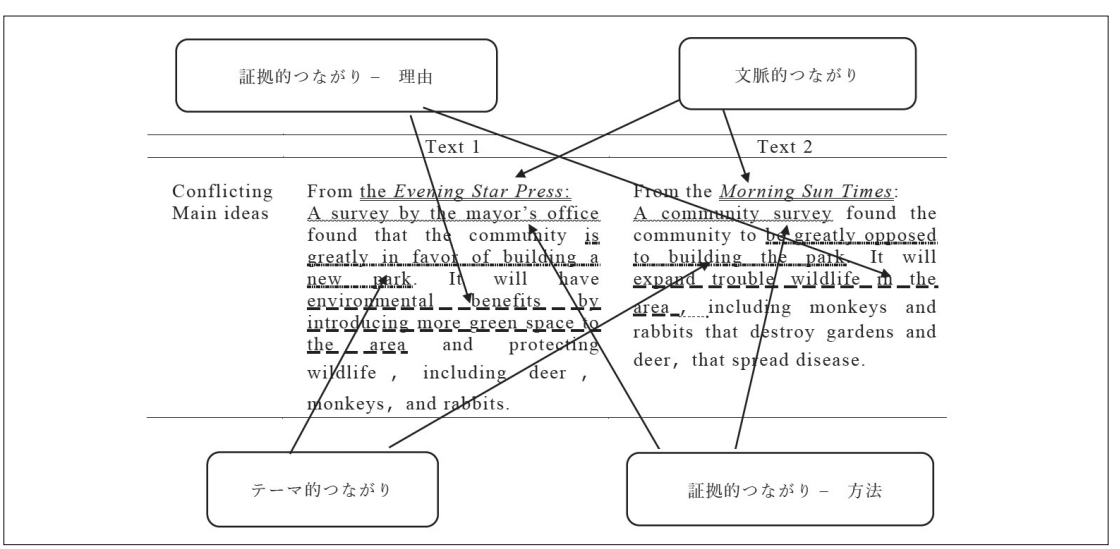
3.1 協力者

大学生, 大学院生32名(男性9名, 女性23名)の協力者内計画で行われた(平均年齢 = 22.69, range = 19-32, *SD*=3.45)。熟達度レベルは英検準1級から2級で, 英検が公表している換算表に基づくとCEFRレベルはB2-B1であった(日本英語検定協会, n.d.)。

3.2 マテリアル

List et al.(2021a)の複数テキストのセットを使用し, 協力者は異なる2つのテキストセットを4つ読んだ。サンプルテキストは以下の通りであった(図2)。個々のテキストは互いに, (1)テキストは新聞記事である点, (2)データの収集方法の情報, (3)中心的ステートメント(住民は新規建造物に賛成または反対)の情報, (4)中心的ステートメントをサポートする根拠情報という4つの要素を含んでいた。2つのテキストセットを読むことで, 読み手は証拠的つながり, テーマ的つながり, そして文脈的つながり(List & Alexander, 2019)の3つの特定性が形成できるよう意図されていた。図2はサンプルテキストセットと形成されるつながりを示している。

2つのテキストセットは4(4つのテキスト条件)×4(4つのテーマ)=16種類用意された(図3)。テキスト条件は次の4つ(相補的主張+重複根拠, 相補的主張+特異根拠, 対立的主張+重複根拠, 対立的主張+特異根拠)であった。主張の関係性は2つであった。相補的主張とは, 住民は施設建設に賛成意見で一致しており, 2つのテキストは補完しあう関係性を持っていた。対立的主張とは, 施設建設への住民意見は賛成と反対で対立しており, 2つのテキストは対立しあう関係性であった。次にテキスト内の根拠内容は2種類あった。テキストセット間で重複するタイプ(重複根拠; e.g., 根拠理由が両方とも商業的視点に基づくものである等)と, 明らかに異なるタイプ(特異根拠;e.g., 根拠理由が商業的視点と環境的視点の明らかに異なるものである等)であった。テーマは公園・高速道路・学校・ショッピングモールの4つについての建設の可否であった。テキスト条件のカウンターバランスを取るため, 16の内4つのテキストセットはランダムに割り当てられた(図4)。表4はテキストの語数とリーダビリティを示している。



■図2: サンプルテキストセットと形成されるつながり

■表4: 各テキストの語数とリーダビリティ

		相補的主張 根拠重複型			相補的主張 根拠特異型			対立的主張 根拠重複型			対立的主張 根拠特異型		
			Text 1	Text 2		Text 1	Text 2		Text 1	Text 2		Text 1	Text 2
School	Words	A1	39	50	A2	39	45	A3	41	41	A4	41	39
	Flesch Reading Ease		65.5	56.2		63.3	54.2		60.1	58.8		56.7	65.5
	FKGL		8.9	11.6		9.2	11.2		9.9	8.4		8.7	8.9
Highway	Words	B1	41	41	B2	42	41	B3	37	41	B4	40	42
	Flesch Reading Ease		58.0	53.9		44.5	58.0		50.8	45.7		57.5	50.5
	FKGL		10.2	10.8		12.2	10.2		10.7	11.9		10.2	11.4
Shopping mall	Words	C1	39	49	C2	39	43	C3	41	41	C4	41	45
	Flesch Reading Ease		63.3	59.3		63.3	49.2		60.1	62.2		60.1	59.9
	FKGL		9.2	11.0		9.2	11.7		9.9	9.6		9.9	10.4
Park	Words	D1	45	43	D2	45	41	D3	41	45	D4	43	45
	Flesch Reading Ease		67.5	57.1		48.6	51.9		55.4	48.6		59.0	48.6
	FKGL		7.5	10.6		12.0	11.1		10.4	12.0		10.3	12.0

注. FKGL= Flesch-Kincaid Grade Level

School	Highway	Shopping Mall	Park
A1 トピック: School 主張: 相補的 根拠: 重複	B1 トピック: Highway 主張: 相補的 根拠: 重複	C1 トピック: Shopping Mall 主張: 相補的 根拠: 重複	D1 トピック: Park 主張: 相補的 根拠: 重複
A2 トピック: School 主張: 相補的 根拠: 特異	B2 トピック: Highway 主張: 相補的 根拠: 特異	C2 トピック: Shopping Mall 主張: 相補的 根拠: 特異	D2 トピック: Park 主張: 相補的 根拠: 特異
A3 トピック: School 主張: 対立的 根拠: 重複	B3 トピック: Highway 主張: 対立的 根拠: 重複	C3 トピック: Shopping Mall 主張: 対立的 根拠: 重複	D3 トピック: Park 主張: 対立的 根拠: 重複
A4 トピック: School 主張: 対立的 根拠: 特異	B4 トピック: Highway 主張: 対立的 根拠: 特異	C4 トピック: Shopping Mall 主張: 対立的 根拠: 特異	D4 トピック: Park 主張: 対立的 根拠: 特異

■図3: テキストセットの条件

条件 1	条件 2	条件 3	条件 4
A1, B2, C3, D4	A1, C4, D3, B2	A1, D4, C2, B3	A1, C4, B3, D2
条件 5	条件 6	条件 7	条件 8
C3, A2, B4, D1	D3, A2, C1, B4	C4, B3, A2, D1	D4, B3, A2, C1
条件 9	条件 10	条件 11	条件 12
C4, D2, A3, B1	B2, D4, C1, A3	B2, C4, D1, A3	D2, B4, C1, A3
条件 13	条件 14	条件 15	条件 16
B3, D2, C1, A4	B2, C3, D1, A4	B2, D3, C1, A4	C3, D2, B1, A4

■図4: 条件間におけるテキストのカウンターバランス

例えば，SchoolはA1からA4まで4種類があり，それぞれに特徴が異なっていた。協力者全員を条件1〜16のどれかにランダムに割り当てた。例えば条件1に割り当てられた協力者は，最初にA1を読んで活動を行い，次にB2を，次にC3を読み，最後にD4を読んだ。

熟達度テスト

すべての協力者は実験後に熟達度テストを受けた。熟達度テストでは，協力者は英検準1級の読解問題16問を15分間で解答するよう指示された。熟達度テストのスコアは15点満点中，M=9.31，SD=2.28，Cronbach's α = 0.79であった。

3.3 手順

本実験は遠隔同時会議サービス（Zoom）を用いた。実験マテリアルはMicrosoft Wordファイルで作成され，協力者は自身のPC上にファイルをダウンロードし，ファイル上で読解とライティングを行った。実験中，協力者のWordファイルは実験実施者と画面共有された。協力者はファイルの画面をスクロールすることなく2つのテキストセットを読むことが出来た。

各テストでの具体的な指示は以下の通りであった（図5）。第一に，協力者には導入的な情報が提示された。その情報とは，「架空の町の市長は新しい施設建設（学校・高速道路・ショッピングモール・公園）を検討している」ことと，「協力者は施設建設についての2つの新聞記事を読んで，その後町の人々が新規施設建設計画についてどのように感じていたかについてライティングすることになる」ことであった。第二に，協力者は相補的または対立的情報かつ重複根拠（根拠の内容が重複するもの）または特異根拠（根拠の内容が異なる観点で構成されているもの）が提示されている2つのテキストを読むよう指示された。そして2つのテキスト間で関係するところをできる限り多く特定することが求められた（約5分）。第三に，特定したテキスト間のつながりを，自身のPC上のWordファイルに日本語で記述することが求められた（約5分）。この時テキストの参照は可とされた。第四に，再度2つのテキストを読んだ協力者は自由回答式の要約設問が提示され，2つのテキストを統合することが求められた（約10分）。町の人々が新規施設建設計画についてどのように感じていたかを50語程度の英語でライティングすることが求められた。テキスト間の全般的な統合的心的表象の把握のため，読後の要約ライティング時にテキストの参照は不可とされた。第五に，協力者はテキスト条件を把握できているかと，2つのテキスト情報を統合的に理解できているかについての3つのYes/No確認問題への回答が求められた（約3分）。スコアは最低が0点，最高が3点と設定された。

■図5：実験手順

確認問題は研究者によって自作された。各テストの前に，練習セッションが実施された。練習セッションでは本実験とは無関係であるが同じ要素を含む2つの新聞記事が用いられたが，その後の本テストと同じ指示が与えられた。協力者からの質疑応答に対応したのち，本テストに進んだ。

3.4 採点

本研究では，(1)読解中に特定されたつながりの形成数と種類(即時処理)，(2)読後の要約ライティング(事後処理)を採点した。

(1)複数テキスト間のつながり形成

複数テキストの読解中に気づいたつながりを日本語で記述させ，即時処理での統合的つながり形成を測定した。分析観点はテキスト間統合の3つの特定性(証拠的，テーマ的，文脈的つながり)を用いた(List et al., 2019)。読解時に形成されたテキスト間のつながりを異なる3つの特定性で分類し，テキスト条件の違いが読解処理に与える効果を検証した。誤って形成されているつながりは採点対象から除外した。複数テキストの読解中に読み手が気づいたつながりについての記述は，先行研究のフレームワークで用いられたコーディングシステム(coding system)に基づいて採点された（List et al., 2021a)。そして証拠的つながり(方法と根拠)，テーマ的つながり，文脈的つながりを反映しているかでカテゴリー化された。

読み手は1つのテキストセットに対し，最大4つのつながりを形成し得た。テキスト間の特定性の種類である2つの証拠的つながり(i.e., データの入手方法と提供された根拠)，1つのテーマ的つながりと，1つの文脈的つながりを潜在的に特定し得るテキスト構造であった。読み手は各テキストセットに対して最大で4つのつながりを特定し得るため，テキスト間で形成された4つの特定性に対して0/1の点数が付与された(採点例は資料1を参照)。

2人の採点者は協力者のプロトコルの30%に対し，個々の反応を上記の3つの領域に分類した。採点者間一致率は高かった(91.7%, Cohen's k = .90)。不一致の部分は2人で話し合って解決し，その点を採点時の注意点としてガイドラインに記載した。それに基づき，残りのデータは著者によって処理された。

(2)要約ライティングの統合についての包括的採点

List et al.（2021a)のルーブリックを用い，メインアイディアが統合されているかを4段階(0〜3点)で採点した(表5)。このルーブリックはつながりの特定性の3種類の観点から，複数テキストの統合的理解を示す統合的な心的表象を把握することを意図していた。ライティングによる要約理解課題ではテキスト参照は不可とした。理由は，要約質問への協力者の反応は統合的な心的表象との一致が意図されるためであった(Britt et al., 1999; Perfetti et al., 1999)。協力者は要約ライティングは英語で書くよう指示された。

先行研究に基づき，List et al.（2021a)の包括的統合スコアルーブリックによる点数付けを行った。採点官は2名の日本人英語教師であった。つながり形成と同様に，2人の採点者は協力者のプロトコルの30%に対し，個々の反応を上記の3つの領域に分類した。採点者間一致率は高かった(83.3%, Cohen's k = .64)。採点例は資料2に示した。

■表5: 統合要約ライティングのルーブリック(List et al, 2021a)

0点	テキスト間のメインアイディアが正確に統合されていない
1点	メインアイディアが正確に統合されている (i.e., 住民は新しい建造物に賛成または反対であることを述べている)
2点	メインアイディアとソースからのサポート情報が正確に統合されている (i.e., 各テキストの具体的な事例が主張のサポートとして用いられている)
3点	メインアイディアが正しく統合され，サポート理由も提供され，この情報を一致する新聞ソースと属性づけている

3.5 統計分析

統計分析1から4はいずれも統計分析ソフトR(version 4.2.2, R Core Team, 2020)を用いた。分析1では,つながり形成が主張・根拠で異なるかを線形混合効果モデル(Linear Mixed Effect Model; LME)によって検討した。従属変数はつながり数であった(資料1の合計参照)。固定効果はカテゴリカル変数である主張要因(相補的と対立的)と根拠要因(重複型と特異型)およびその交互作用であった。交互作用を含むために,これらのカテゴリカル変数はsum-codingとしてRで読み込まれた。共変量として,熟達度テストのスコアをz変換した上でモデルに組み込んだ。熟達度テストのスコアにランダム傾きとして協力者を加えたモデルを作成した。LMEは説明変数の欠損値に頑健な分析モデルであり,ランダム傾きとして協力者の個人差を考慮できることから分散分析よりも詳細に結果を分析できる。モデルの作成にはlme4パッケージ(version 1.1-32; Bates et al., 2015),下位検定にはemmeansパッケージ (version 1.8.5; Lenth, 2023),プロットにはsjPlotパッケージ(version 2.8.13; Lüdecke, 2023)を用いた。

分析2では,異なる特定性のつながり形成数を主張要因(相補的と対立的)と根拠要因(重複根拠と特異根拠)ごとに合計し従属変数とし,主張要因と根拠要因で異なるかをウィルコクソンの符号付順位検定(Wilcoxon-signed rank test)で分析した。ウィルコクソンの符号付順位検定は対応のある*t*検定に相当し,比較する母集団に確率分布を前提としないため,頻度などの名義尺度データに対しても使用できる(平井他, 2022)。分析手法はList et al.(2021a)に準じた。例えば,テーマ的つながりの場合,主張要因を次のように集計した(資料1の主張要因別スコアを参照)。相補的＋重複根拠の点数(0-1点)と,相補的＋相違根拠の点数(0-1点)を加算し,相補的なテーマ的つながりの合計(0-2点)を算出した。また対立的＋重複根拠の点数(0-1点)と,対立的＋相違根拠の点数(0-1点)を加算し,対立的なテーマ的つながりの合計(0-2点)を算出し,この2つを比較した。つぎに根拠要因は次のように集計した(資料1の根拠要因別スコアを参照)。相補的＋重複根拠の点数(0-1点)と,対立的＋重複根拠の点数(0-1点)を加算し,重複根拠の合計(0-2点)を算出した。また相補的＋相違根拠の点数(0-1点)と,対立的＋相違根拠の点数(0-1点)を加算し,相違根拠の合計(0-2点)を算出し,この2つを比較した。

分析3の,統合的理解の分析は,要約ライティングスコアを従属変数とし,主張要因(相補的と対立的)×根拠要因(重複根拠と特異根拠)を独立変数とする2×2の二元配置分散分析(repeated ANOVA)で分析した。Mauchlyの球面性検定を実施したところ,$\chi^2(5) = 2.412$, $p = .789$と有意ではなかった。球面性の仮定が成立しており,対応あり要因の水準間の差の分散が等しく,ANOVA実施の前提を満たしていた。

分析4の,つながり形成と統合的理解の関係についての分析は,総つながり形成数・特定性別のつながり形成数(証拠的つながり,テーマ的つながり,文脈的つながり),要約ライティングスコアの相関分析をピアソンの相関係数を用いて実施した。

4 結果

4.1 記述統計

表6は総つながり数の記述統計を示している(相補的主張: $M = 1.81$, $SD = 1.10$; 対立的主張: $M = 1.91$, $SD = 0.93$; 重複根拠: $M = 1.83$, $SD = 0.97$; 特異根拠: $M = 1.89$, $SD = 1.05$)。表7は証拠的つながり,表8はテーマ的つながりと文脈的つながりの記述統計を示している。相補的テキストと対立的テキストのいずれにおいても,平均的な読み手が形成したつながりは,証拠的つながり(理由),証拠的つながり(方法),テーマ的つながり,そして文脈的つながりの合計である最大形成可能数(4つのつながり)の半分以下であった。

■表6: 総つながり数と要約ライティングスコア

	総つながり数					要約ライティングスコア				
	M	*SD*	Skewness	Kurtosis	*SE*	*M*	*SD*	Skewness	Kurtosis	*SE*
①相補的＋重複根拠	1.72	1.11	-0.13	-0.97	0.20	1.56	0.84	-0.03	-0.70	0.15
②相補的＋特異根拠	1.91	1.20	-0.26	-1.00	0.21	1.56	1.05	-0.24	-1.21	0.18
③対立的＋重複根拠	1.94	1.08	-0.33	-0.86	0.19	1.59	0.98	-0.15	-1.06	0.17
④対立的＋特異根拠	1.88	1.07	-0.22	-0.89	0.19	1.47	0.98	-0.21	-1.13	0.17
⑤相補的主張(①+②)	1.81	1.10	-0.14	-0.93	0.19	1.56	0.87	-0.19	-0.92	0.15
⑥対立的主張(③+④)	1.91	0.93	-0.50	-0.70	0.16	1.53	0.89	-0.12	-0.76	0.16
⑦重複根拠(①+③)	1.83	0.97	-0.37	-0.90	0.17	1.58	0.81	-0.11	-0.56	0.14
⑧特異根拠(②+④)	1.89	1.05	-0.22	-1.12	0.19	1.52	0.93	-0.07	-1.05	0.16
⑨合計	1.86	0.96	-0.39	-0.92	0.17	1.55	0.83	-0.09	-0.63	0.15
⑩低熟達成度	1.64	1.05	-0.23	-1.35	0.26	1.27	0.82	-0.34	-1.48	0.21
⑪高熟達成度	2.08	0.84	-0.29	-1.05	0.21	1.83	0.76	0.40	-1.30	0.19

注1. 総つながり数は4点満点, 要約ライティングは3点満点である
注2. 熟達度テストのスコアを基に, 協力者を低熟達度と高熟達度に分類された

■表7: 証拠的つながり-根拠と証拠的つながり-方法

	証拠的つながり - 根拠					証拠的つながり - 方法				
	M	*SD*	Skewness	Kurtosis	*SE*	*M*	*SD*	Skewness	Kurtosis	*SE*
①相補的＋重複根拠	0.59	0.50	-0.36	-1.92	0.09	0.12	0.34	2.16	2.76	0.06
②相補的＋特異根拠	0.66	0.48	-0.63	-1.66	0.09	0.12	034	2.16	2.76	0.06
③対立的＋重複根拠	0.62	0.49	-0.49	-1.81	0.09	0.12	034	2.16	2.76	0.06
④対立的＋特異根拠	0.44	0.50	0.24	-2.00	0.09	0.16	037	1.81	1.30	0.07
⑤相補的主張(①+②)	0.62	0.42	-0.47	-1.47	0.07	0.12	0.25	1.79	2.34	0.04
⑥対立的主張(③+④)	0.53	0.44	-0.11	-1.73	0.08	0.14	0.26	1.56	1.50	0.05
⑦重複根拠(①+③)	0.61	0.40	-0.38	-1.36	0.07	0.12	0.25	1.79	2.34	0.04
⑧特異根拠(②+④)	0.55	0.41	-0.16	-1.53	0.07	0.14	0.32	1.93	2.24	0.06
⑨合計	0.58	0.37	-0.23	-1.39	0.06	0.13	0.21	1.32	0.59	0.04
⑩低熟達成度	0.55	0.43	-0.12	-1.84	0.11	0.09	0.20	2.17	4.13	0.05
⑪高熟達成度	0.61	0.30	-0.18	-1.01	0.08	0.17	0.22	0.58	-1.50	0.05

注. 証拠的つながり - 根拠と証拠的つながり - 方法は各4点満点である

■表8: テーマ的つながりと文脈的つながり

	テーマ的つながり					文脈的つながり				
	M	*SD*	Skewness	Kurtosis	*SE*	*M*	*SD*	Skewness	Kurtosis	*SE*
①相補的＋重複根拠	0.59	0.50	-0.36	-1.92	0.09	0.41	0.50	0.36	-1.92	0.09
②相補的＋特異根拠	0.69	0.47	-0.77	-1.45	0.08	0.44	0.50	0.24	-2.00	0.09
③対立的＋重複根拠	0.81	0.40	-1.53	0.34	0.07	0.38	0.49	0.49	-1.81	0.09
④対立的＋特異根拠	0.84	0.37	-1.81	-1.30	0.07	0.44	0.50	0.24	-2.00	0.09
⑤相補的主張(①+②)	0.64	0.41	-0.52	-1.34	0.07	0.42	0.48	0.30	-1.87	0.08
⑥対立的主張(③+④)	0.83	0.30	-1.45	0.98	0.05	0.41	0.43	0.35	-1.60	0.08
⑦重複根拠(①+③)	0.70	0.38	-0.78	-0.89	0.07	0.39	0.45	0.42	-1.68	0.08
⑧特異根拠(②+④)	0.70	0.38	-0.78	-0.89	0.07	0.44	0.49	0.24	-1.94	0.09
⑨合計	0.73	0.03	-1.00	0.08	0.05	0.41	0.44	0.31	-1.75	0.08
⑩低熟達成度	0.67	0.35	-0.70	-0.87	0.09	0.33	0.42	0.59	-1.43	0.10
⑪高熟達成度	0.80	0.23	-0.84	-0.28	0.06	0.50	0.47	0.00	-1.98	0.12

注. テーマ的つながりと文脈的つながりは各4点満点である

表9は各テストの最後に実施する確認問題の結果である。全体でのスコアは*M*=2.15, *SD*=0.33, *Min*=0, *Max*=3であった。テキスト条件ごとのスコアを見ると, 相補的重複が最もスコアが高かった(*M*=2.31, *SD*=0.59)。一方最も低かったテキスト条件は対立的重複であった(*M*=2.00, *SD*=0.67)。最大可能スコアが3点であることを踏まえると, 協力者にとって本テキストセットを完全に理解できていない可能性が示唆された。

■表9: 確認問題のスコア

	M	*SD*	Skewness	Kurtosis	*SE*
相補的+重複根拠	2.31	0.59	-0.18	-0.77	0.10
相補的+特異根拠	2.12	0.66	-0.12	-0.81	0.12
対立的+重複根拠	2.00	0.67	0.00	-0.85	0.12
対立的+特異根拠	2.16	0.77	-0.67	0.11	0.14
合計	2.15	0.33	-0.73	-0.10	0.06

4.2 つながり形成

4.2.1 総つながり数

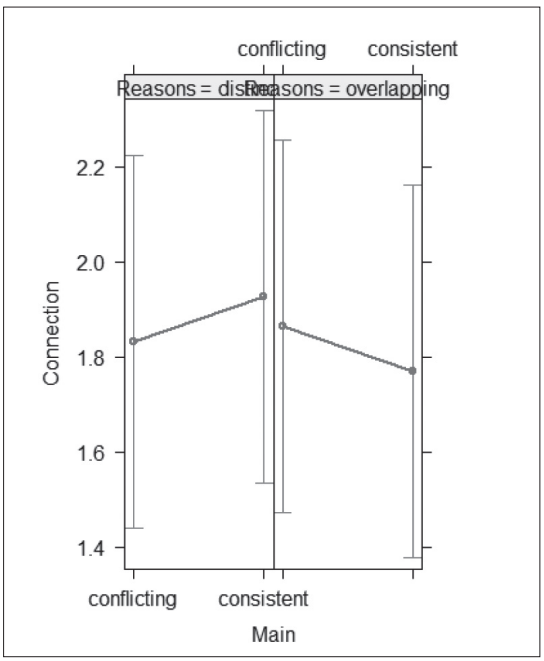
表10は分析1の結果である。LMEの結果, 熟達度の主効果が有意であった（β = .39, *SE* = 0.11, z = 3.45, p = .002)。このことから, 熟達度の高い学習者ほど, つながり形成を行う傾向がある程度あることが示された。共変量である熟達度の影響を取り除いた場合の主張要因と根拠要因の交互作用は非有意であった（β = −.05, *SE* = 0.06, z = −0.80, p = .426)。また主張要因(β = −.00, *SE* = 0.06, z = −0.00, p = 1.000), 根拠要因の主効果も非有意であった（β = .03, *SE* = 0.06, z = 0.53, p = .595)。つまり主張要因, 根拠要因ともに総つながり数に有意な差はなく, これらのテキストに関する要因からの影響は見られなかった(図6)。

■表10: LME分析の結果

Predictors	Estimates	95%CI	*t* value	*p* value
(Intercept)	1.85	1.51–2.18	10.87	<.001
Main	0.00	-0.12–0.12	0.00	1.000
Reasons	0.03	-0.08–0.15	0.53	.595
Proficiency	0.39	0.17–0.61	3.45	.001
Main × Reasons	-0.05	-0.16–0.07	-0.80	.426
Random Effects				
σ^2	0.44			
τ 00 participant	0.76			
τ 11 participant proficiency	0.08			
ρ 01 participant	-1			
N participant	32			
Observations	128			
Marginal R^2/Conditional R^2	0.261/NA			

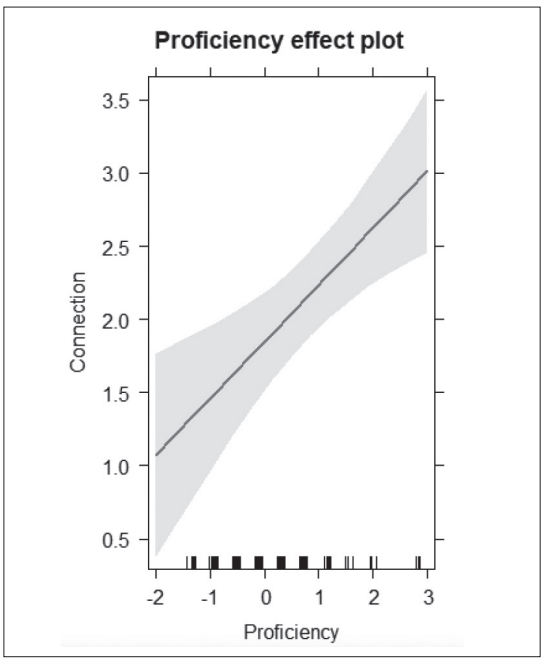
注. モデルは以下の通りである:
Connection ～ Main * Reasons + Proficiency + (1 + Proficiency | Participant)

熟達度とつながり数をプロットした図7では，x軸にはz変換された熟達度のスコアが，y軸にはつながり数がプロットされている。共変量としての熟達度要因が有意であったことは，x軸の値が大きくなり右に向かうほどy軸の値が上昇する右肩上がりのグラフであることに表されている。熟達度が高い人はつながり形成数が多く，熟達度要因はつながり形成に影響していた。



■図6：テキスト関係に対する主張要因と根拠要因のプロット図

注．左ボックスの左は対立的＋特異根拠（つながり形成数の平均値と*SD*），左ボックスの右は対立的＋重複根拠，右ボックスの左は相補的＋特異根拠，右ボックスの右は相補的＋重複根拠を示している



■図7：熟達度とつながり数のプロット図

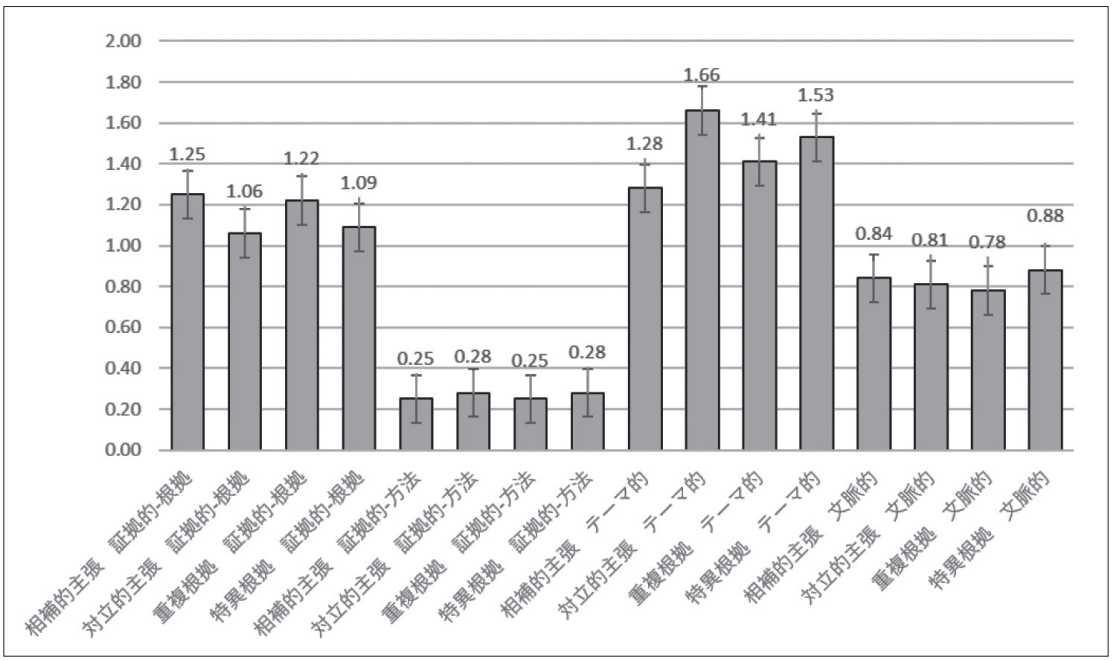
注．x軸は熟達度（0を平均としてセンタリングした値），y軸はつながり形成数（平均値と95%信頼区間）を示している

4.2.2 各つながりの形成数と種類

図8と表11は分析2の記述統計を示している。分析2では，3つの特定性（証拠的つながり－方法，証拠的つながり－根拠，テーマ的つながり，文脈的つながり）におけるつながり数を，Wilcoxon-signed rank testを用い，異なる4つのテキスト条件（対立的主張，相補的主張，重複根拠，特異根拠）でそれぞれ比較した（図8）。

各つながりの形成数と種類についてはテーマ的つながりにおいてのみ，対立的テキスト（M = 1.66, SD = 0.60）に対し，相補的テキスト（M = 1.28, SD = 0.81）で有意差があり（Z = 2.36, p = .018），テーマ的つながりの主張間条件にのみ有意差が見られた。

しかし他のテキスト要因による違いは見られなかった。テーマ的つながりの重複的テキスト（M = 1.41, SD = 0.80）に対し，特異的テキスト（M = 1.53, SD = 0.72）で非有意（Z =0.88, p = .381）であった。証拠的つながり－根拠は，対立的テキスト（M = 1.06, SD = 0.88）に対し，相補的テキスト（M = 1.25, SD = 0.84）で非有意（Z =1.29, p = .197）であり，重複的テキスト（M = 1.22, SD = 0.79）に対し，特異的テキスト（M = 1.09, SD = 0.82）で非有意（Z =0.79, p = .431）であった。証拠的つながり－方法は，対立的テキスト（M = 0.28, SD = 0.52）に対し，相補的テキスト（M = 0.25, SD = 0.51）で非有意（Z=0.00, p = 1.000）であり，重複的テキスト（M = 0.28, SD = 0.63）に対し，特異的テキスト（M = 0.25, SD = 0.51）で非有意（Z =0.16, p = .875）であった。文脈的つながりは，対立的テキスト（M = 0.81, SD = 0.86）に対し，相補的テキスト（M = 0.84, SD = 0.95）で非有意（Z =0.00, p = 1.000）であり，重複的テキスト（M = 0.78, SD = 0.91）に対し，特異的テキスト（M = 0.88, SD = 0.98）で非有意（Z=0.78, p = .438）であった。



■図8: 主張要因と根拠要因のつながり形成数

注. ひげの最大値と最小値は標準誤差(*SE*)で, 各つながりの最大値は2である

■表11: 相補的主張・対立的主張・重複根拠・特異根拠でのつながり数

	証拠的つながり-根拠					証拠的つながり-方法				
	M	*SD*	Skewness	Kurtosis	*SE*	*M*	*SD*	Skewness	Kurtosis	*SE*
相補的主張	1.25	0.84	-0.47	-1.47	0.15	0.25	0.51	1.79	2.34	0.09
対立的主張	1.06	0.88	-0.11	-1.73	0.16	0.28	0.52	1.56	1.50	0.09
重複根拠	1.22	0.79	-0.38	-1.36	0.14	0.25	0.51	1.79	2.34	0.09
特異根拠	1.09	0.82	0.82	-1.53	0.14	0.28	0.63	1.93	2.24	0.11

	テーマ的つながり					文脈的つながり				
	M	*SD*	Skewness	Kurtosis	*SE*	*M*	*SD*	Skewness	Kurtosis	*SE*
相補的主張	1.28	0.81	-0.52	-1.34	0.14	0.84	0.95	0.30	-1.87	0.17
対立的主張	1.66	0.60	-1.45	0.98	0.11	0.81	0.86	0.35	-1.60	0.15
重複根拠	1.41	0.76	-0.78	-0.89	0.13	0.78	0.91	0.42	-1.68	0.16
特異根拠	1.53	0.67	-1.04	-0.20	0.12	0.88	0.98	0.24	-1.94	0.17

注. 最大値は2である(e.g., 相補的主張のつながり形成可能数は2であるが, その内訳は相補的重複根拠の1, 相補的特異根拠の1を合算したものである)

4.3 要約ライティングスコア

分析3の記述統計は表6で示している。主張要因, 根拠要因をrepeated ANOVAで比較した結果, 主張要因[$F(1, 31) = 0.092$, $p = .763$, $\eta^2 < .001$], 根拠要因[$F(1, 31) = 0.413$, $p = .525$, $\eta^2 < .001$]で条件間の要約スコアに差はなかった。また交互作用もみられなかった。つまり複数テキスト間の関係性は要約ライティングに影響を与えなかった(表12)。

■表12: 要約ライティングスコアのANOVAテーブル

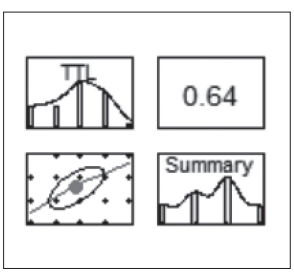
Predictors	*SS*	*df*	*MS*	*F*-value	*p*-value	η^2
main_ideas	0.031	1	0.031	0.093	.763	0.001
s x Main_ideas	10.469	31	0.388			
overlapping_distinct	0.125	1	0.125	0.413	.525	0.001
s x overlapping_distinct	9.375	31	0.302			
main_ideas x overlapping_distinct	0.125	1	0.125	0.374	.546	0.001
s x Main_ideas x overlapping_distinct	10.375	31	0.335			
Total	115.719	127	0.911			

4.4 相関分析

分析4において, 形成された全つながりの総数と要約ライティングスコアの相関分析を行ったところ, r (32)= .64, p < .001, 95%CI (0.53, 0.74)と中程度の相関があり, 0.1%水準で有意であった(図9)。

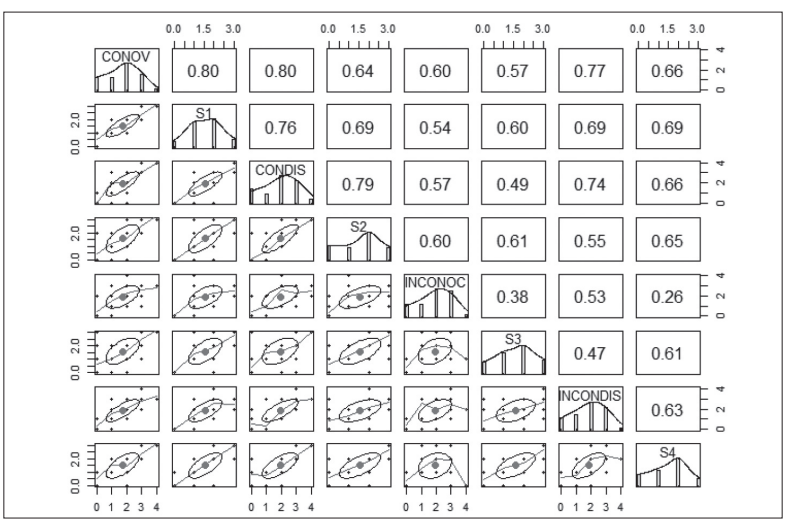
個々のつながり形成別の相関関係は以下の通りである(図10参照。その記述統計は表6参照)。個別に形成されたつながりと要約ライティングスコアにおいても相関関係がみられた。形成された相補的主張かつ根拠重複型の総数と要約ライティングスコア間には, r (32)= .80, p < .001, 95%CI (0.62, 0.90)と強い相関があり, 0.1%水準で有意であった。形成された相補的主張かつ根拠特異型の総数と要約ライティングスコア間には, r (32)= .79, p < .001, 95%CI (0.61, 0.89)と強い相関があり, 0.1%水準で有意であった。

また形成された対立的主張かつ根拠重複型の総数と要約ライティングスコア間には, r (32)= .38, p < .05, 95%CI (0.03, 0.64)と弱い相関があり, 5%水準で有意であった。形成された対立的主張かつ根拠特異型の総数と要約ライティングスコア間には, r (32)= .63, p < .001, 95%CI (0.35, 0.80)と中程度の相関があり, 0.1%水準で有意であった。つまり複数テキストを読解した際のテキスト間のつながり形成と読後の統合的理解には関連性が見られたが, 相補的テキストのほうがつながり数と要約スコアの間に高相関がみられた一方, 対立的テキストは弱から中程度の相関にとどまった。



■図9: 総つながり数と要約ライティングスコアの相関図

注. TTL=総つながり数; Summary=要約ライティングスコア



■図10: テキスト条件別のつながり形成数と要約ライティングスコアの相関関係図

注. CONOV=相補的重複; CONDIS=相補的特異; INCONOC=対立的重複; INCONDIS=対立的特異;S1=相補的重複の要約ライティングスコア; S2=相補的特異の要約ライティングスコア; S3=対立的重複の要約ライティングスコア; S4=対立的特異の要約ライティングスコア

5 考察

RQ1 **テキスト間の関係性を統制した複数の説明文テキストを読解するとき，形成されるテキスト間のつながり総数，異なる特定性のつながり数はテキスト要因において違いがみられるか。**

5.1 総つながり数（分析1）

総つながり数に関しては，三つの予想を設定した。予想1では，L2学習者は対立的テキストの処理に困難を抱えるため，総つながり数はテキストの関係性によって影響を受け，特に対立的テキストでつながり形成数が少ないことを想定していた。予想2では対立的テキストでつながり形成数が多いと予想した。予想3では，総つながり数はテキスト関係によって影響を受けないことを想定していた。分析1の結果，主張要因，根拠要因ともに有意差はなく，テキスト要因での影響は見られなかった（予想3を支持）。この結果は，L1の先行研究と異なるものであった。List et al.（2021a）によると，相補的テキストのほうが総つながり数は多かったが，今回のEFL学習者を対象とした実験では，総つながり数に対するテキスト要因の影響は見られなかった。この理由の一つは，L1話者において想定されている対立的テキスト処理のモデルが，L2学習者にはうまく機能しないことが考えられる。

今回の結果は，テキストの関係性に焦点を当てた読解モデルであるD-ISCモデルによって説明しうる。D-ISCモデルでは，相補的テキストと対立的テキストでは読み手の認知処理が異なることを仮定している（Braasch & Kessler, 2021）。相補的テキストでは，次の処理が仮定される。相補的テキストの場合，情報ソース間で高い一致がみられた内容は，長期的記憶から活性化された情報と容易に一致するため，情報の整合性は容易に獲得される（van den Broek et al., 1995）。認知的不均衡を生じないため，読み手はソースの精緻化処理への認知努力を高めない。一方，対立的テキストでは，テキスト内でのくいちがいを認知すると，熟達した読み手はゆっくり読む（e.g., Beker et al., 2016）。その結果，認知的均衡の再構築が刺激されると言われている。このモデルはL1話者を想定しているが，L2学習者ではこの読解モデルの核である認知的不均衡を経験しない可能性がある。つまりL2学習者はくいちがいに気づかない傾向があり，理解の破綻なしに，ソースの特徴への注意を評価せず，読み手は読解を続けることがある。（Braasch & Kessler, 2021）。今回の結果で総つながり形成数にテキスト間で差が見られなかったのは，このようなL2学習者の特性が影響した可能性がある。

一方，熟達度の主効果が有意で熟達度が高い人はつながり形成数が多い傾向にあり，熟達度要因はつながり形成に影響していた。当初の予想1と一致するものであるが，先行研究を深く読むと，次のような理由が考えられる。理由の一つは，L2では高熟達度学習者のほうがグローバルな読解方略を使用する傾向が高いことが考えられる。L2研究では，複数テキストの読解中，対立的テキストを高熟達学習者と低熟達学習者が読解した場合，高熟達度学習者のほうが複数テキストの内容全体を把握しようと努めるのに対し，低熟達度学習者は単語レベルの理解に固執していた（Plakans, 2009）。特に複数テキストを理解する場合，個々のテキストは物理的に離れており，またほかのテキストとの明示的な関係に言及していないため，テキスト間の統合またはつながりを形成することはL1の読み手にとっても難しい（List, 2020）。L2学習者にはなおさら困難であることが想定されていたが，今回の結果は学習者の熟達度によってテキスト間のつながり形成に差がみられた。つまりこの結果は，複数テキストの内容全体を把握しようと努める高熟達度学習者の読解方略の巧拙が，つながり形成の鍵であることを示唆している。

5.2 つながりの特定性（分析２）

表13はL1研究と本研究における特定性別のつながり形成のまとめである。特定性別のつながり形成数については，事前の予想では，トレードオフ仮説と，List et al.（2021a）に基づき，文脈的つながりは対立

的テキストの方で多く形成される一方，証拠的つながりとテーマ的つながりは相補的テキストでより多く形成されること（予想1）と，テキストの関係性は，つながりの特定性に影響を与えない（予想2）という2つを想定していた。本実験では，つながり形成数に差が見られたのは，テーマ的つながりのみで，対立テキストのほうがつながり形成数は有意に多かった（予想2に該当）。

■表13: 特定性別のつながり形成のまとめ

	L1（List et al., 2021a）		L2（本研究）	
	相補的	対立的	相補的	対立的
総つながり数	●	＞		
証拠的つながり	●	＞		
テーマ的つながり	●	＞	＜	●
文脈的つながり	＜	●		

考えられる理由としては，日本人学習者は複数テキストの関係性を整理してとらえることが苦手な可能性がある。トレードオフ仮説は，読み手がテキスト間の関係性を整理して把握できるということを仮定している。例えば相補的テキストは意味の重複が多く，内容のつながりが強化されて自動的に統合が生じる一方，対立的テキストはくいちがう情報によって読み手は認知的不均衡を経験し，ソース情報への注意が促される（Braasch et al., 2016）。確かに先行研究ではL1話者はテキスト関係を捉え，トレードオフ仮説が実証されていた（List et al., 2021）。しかし今回の結果は，EFL学習者が形成したつながり数は形成可能数の半分以下であり，テキストの関係性間での違いも見られず，トレードオフ仮説は実証されなかった。つまり相補的テキストと対立的テキストを読解するとき，EFL学習者は，統合的に理解するために必要な認知的処理を適切に切り替えられていない可能性を示唆している。

今回の結果は，想定していた予想1と異なり，テーマ的つながりのみ，対立テキストのほうがつながり形成数は有意に多かった（予想2）が，これはL1研究のD-ISC仮説はEFL学習者には当てはまらない可能性があることを示唆している。理由は，以下の通りである。

第一に，L1研究では，読み手が対立的テキストを読解するとき，内容のくいちがいを認知する。そしてソース情報などの文脈的つながりを注視することにより，くいちがう内容の均衡回復を図ろうとする（Braasch & Kessler, 2021）。しかしL2学習者の場合，対立情報に出会ったとき，文脈情報を用いて，内容のくいちがいから生じる不均衡を十分認知できておらず，このD-ISCモデルの仮定が機能しない可能性がある。

第二に，L2学習者は2つのテキストの関係性を整理し，一貫した理解を構築することが難しい可能性がある。低熟達度のL2学習者は，局所的な理解に固執する（Plakans, 2009）。このため対立的テキスト読解時，目に留まりやすい対立する情報に注意が向けられてしまい，主題の対立といった把握しやすい表面的な情報に注意がひきずられやすく，根拠や文脈情報へ認知資源が十分に割かれない可能性がある。またEFL学習者は対立テキストを読解しても，ソース情報のようなメタ的文脈情報を活用して，くいちがいを解消しない，またはできない可能性がある。テーマ的つながりは最も形成されやすいことに加え，対立的テキストの場合はその違いが際立つため，EFL学習者にとってつながりを形成しやすかったといえる。

RQ2　テキスト間の関係性を統制した複数の説明文テキストを読解するとき，読解中に形成されるテキスト間のつながりは読後の要約ライティングにおける統合的理解と関係があるか。

5.3 テキスト関係と要約ライティングの関係（分析３）

分析3の結果，複数テキスト間の関係性は要約ライティングのスコアに影響を与えず，予測2に該当した。この結果はテキスト関係が要約ライティングスコアに影響するという先行研究と異なるものであった。

List et al.(2021a) によると, L1話者において, 相補的テキストのほうが対立的テキストよりも要約ライティングスコアが高く, テキスト間の関係性によって要約ライティングのスコアは異なっていた(予測1を支持)。つまりテキストの関係性が要約ライティングに与える影響は, L1とL2で異なっていた。

今回の結果は, テキスト参照を不可とした実験方法がEFL学習者のライティング結果に影響を与えた可能性がある。これまでのL2研究においてテキストの関係性がライティングのスコアに影響した研究では, ライティング時にテキスト参照を許可している。例えばQin and Liu (2021)では, 複数テキストを読解したEFL学習者は, 対立的テキストのほうがライティングスコアは高かったが, ライティング時のテキスト参照は可能であった。しかし本研究では, 先行研究との比較可能性の観点から, 複数テキストを読解した後にテキストの参照を不可とした。要約ライティング時のテキスト参照が認められない条件下では, EFL学習者にとってライティングの認知的負担が高いため, ライティング前のテキスト情報の字義的な暗記を協力者に強いたかもしれない。つまり, 協力者はテキスト情報を記憶することに注力したため, テキストの関係性を丁寧に読んでライティングに反映することが難しかった可能性がある。結果として, L1と異なり, EFL学習者は関係性の異なるテキストを読んでも, ライティング上での違いは見られなかった。

次にL1研究のライティングと異なる結果になった理由は, 協力者が頭の中で理解できていた複数テキストの統合状態が, 英語でのライティングの段階になると適切に産出できなかったことが考えられる。そもそも要約ライティングはL2熟達度の影響が大きい(Johns & Mayes, 1990)。要約ライティングは3点満点中, 平均スコアは1.55と半分程度にとどまっていた。結果として, EFL学習者は複数テキストの異なるテキスト関係を, 英語での要約ライティングに反映できなかったのかもしれない。

5.4 テキスト間のつながりと統合的理解の関係(分析4)

分析4については, 形成されたつながりと要約ライティングスコアにおいて相関関係が見られた。つまり複数テキストの読解中, つながり形成と読後の統合的理解には関連性が見られ, 事前の予測1とある程度一致した。またテキスト関係によってつながり数と要約スコアの相関に差がみられた。相補的テキストのほうがつながり形成と要約スコアの相関が高い一方, 対立テキストは相関が弱から中程度にとどまった。このことから, 対立テキストの理解には, つながり形成数以外の要因もより多く考えられる可能性がある。具体的にみると, 対立的主張かつ根拠重複型のつながり総数と要約ライティングスコア間には, r (32) = .38, p < .05, 95%CI (0.03, 0.64)と, 他のテキスト関係条件と比べて相関が弱かった。これは, つながり形成数と統合的理解には相関がみられる一方で, 対立的テキストの統合的理解を把握するためには, 単に形成されたつながり総数を測定するだけは不十分な可能性があることを示している。

対立的テキストにおいて相関関係が弱かった理由は, 主張の対立しているテキスト読解をする時, 形成されるつながりの数が多くても, 統合的な理解には至らなかったことが想定される。つながり形成には統合の深さがあり, 進行性のレベルが仮定され, 単につながりに気づいているような低次元のレベルから, 内容を精緻に捉えて推論しながらつながりを形成する高次元のレベルが存在する(List et al., 2021b)。今回の実験で対立的テキストでのつながり形成数と要約ライティングの相関が弱かったのは, 形成されたつながりが低次元のレベルにとどまっていた可能性がある。つまりテキスト関係によって生じるつながり形成数だけでは, 統合的理解を十分には把握できないため, つながり数だけでなく, 統合レベルの高低という視点からも検証する必要がある。またテキスト間のどこがつながっているかだけでなく, なぜつながっているのかという因果関係などの視点も含めると, 形成されたつながりをより多面的に検証できる可能性がある。

今回顕著であったのは, 先行研究と比べ, 本研究の相関係数は総じて高かった点である。考えられる理由は2つある。第一に, EFL学習者特有の要因と実験方法の影響が考えられる。Keck(2014)では, L1とL2学生に1000ワードのテキストを用いてテキスト参照可で要約ライティングを課したところ, L2学生のほうがテキスト情報をよりコピーした。このようにL2学習者は要約ライティング時にテキスト情報をそのまま使用する傾向がある。テキストの参照を不可とした今回の実験方法は, EFL学習者にとって負担が重

いため，協力者は字義通りにテキストを暗記し，ライティングに使用した可能性がある。結果として，テキスト情報をそのまま暗記してつながりを表現し，要約ライティングに反映できた学習者は，つながり数と要約ライティングスコアがより高くなった。これにより形成されたつながりと要約ライティングスコアにおいて高い相関関係が見られた可能性がある。

相関関係が高かった第二の理由は，熟達度の問題である。本研究の分析1では，熟達度の高い人はつながり形成数が多かった。特に要約ライティングはL2熟達度の影響が大きい（Johns & Mayes, 1990）。このため高熟達度の学習者のほうが頭の中でおこなっていた複数テキストの統合を英語で適切に産出でき，ライティングにも適切に英語で反映できたので統合スコアが高くなった。一方，低熟達度学習者には逆のことが生じた。複数テキストの読解に苦労するので，特定できたつながり形成数も少なく，英語の産出にも課題があるためライティングにも統合を反映できなかった。結果として，EFL学習者のつながり形成数と要約スコアに高い相関が見られた可能性がある。

6 結論

6.1 結論

本研究の目的は，複数テキストを読解した際のテキスト間のつながり形成と，読後の統合的理解の関係を，即時処理と事後処理の2つの測定方法で調べることである。総つながり数に関しては，主張要因と根拠要因ともに差はなく，テキスト要因での影響は見られなかった。この理由の一つは，L1話者において想定されている対立的テキスト処理のモデルが，L2学習者にはうまく機能しないことが考えられる。一方，熟達度の主効果が有意で熟達度が高い人はつながり形成数が多く，熟達度要因はつながり形成に影響していた。

特定性別のつながり形成数については，トレードオフ仮説と，List et al.（2021a）に基づき，文脈的つながりは対立的テキストの方で多く形成される一方，証拠的つながりとテーマ的つながりは相補的テキストでより多く形成されることを想定していた。しかし本実験では，つながり形成数に差が見られたのは，テーマ的つながりのみで，対立テキストのほうがつながり形成数は有意に多かった。考えられる理由としては以下の通りである。第一に，EFL学習者は複数テキストの関係性を整理してとらえることが苦手な可能性がある。第二に，EFL学習者の場合，対立情報に出会ったとき，内容のくいちがいから生じる不均衡を十分認知できておらず，D-ISCモデルの仮定が機能しない可能性がある。

つながり形成と統合的理解の関係については，個別に形成されたつながりと要約ライティングスコアにおいて，相関関係が見られた。つまり複数テキストを読解した際のテキスト間のつながり形成と読後の統合的理解には関連性が見られた。しかし形成された対立的主張かつ根拠重複型の総数と要約ライティングスコア間の関係は，他のテキスト関係条件と比べて相関が弱かった。これは，つながり形成数と統合的理解には相関がみられる一方で，対立的テキストの統合的理解を把握するためには，単に形成されたつながり総数を測定するだけは不十分な可能性があるからである。特に今回の実験では対立的テキストで形成されたつながりが低次元のレベルにとどまっていた可能性がある。その結果，要約ライティングには高いレベルの統合が表れておらず，低い点数になったと考えられる。

6.2 限界点

本研究の限界点は次の通りである。第一に，本研究と先行研究の結果の相違は，実験方法の違いが結果に影響した可能性がある。本研究は新型コロナウイルス感染症（COVID-19）の影響化において，オンラインで実験を行い，先行研究における対面での実施とは異なっていた。対面での紙を用いた実験に比べ，

デジタル上のマテリアルを読解した場合，テキストの理解度が下がるとされる（Delgado & Salmerón, 2022）。また読み手のメタ的認知行動である，認知している読解行動と，実際の読解行動がどの程度一致しているかという「読み手の調整度（reader's calibration）」は，紙のほうが良く，デジタルのほうが劣る（Clinton, 2019）。EFL学習者の読解処理を考慮すると，オンライン上で複数テキストを読解した本実験は，読み手にとって負荷が高かったため，複数テキストの関係性を整理して理解することが難しかった可能性がある。

第二に，実験のタスク中に与えた使用言語の指示が，今回の結果に影響した可能性がある。テキストを読解している時に特定したつながりは，日本語で記述するよう指示を与えた。一方，要約ライティングは英語で作成するよう指示を与えた。つまり，つながり特定と要約設問の言語が異なっており，協力者が頭の中で形成した複数テキストの統合的な理解が，英語でのライティングに反映できなかった可能性がある。例えば2つのタスクを日本語のみ，または英語のみで行うよう指示した場合，異なる結果が生じるか調査する必要がある。

第三に，実験協力者のモチベーションが実験結果に影響した可能性がある。本研究では大学のジョブ機会プラットフォームを経由し，大学生の実験協力者をリクルートした。一方で先行研究（List et al., 2021a）では，大学の授業を履修する大学生を対象としており，協力者には追加の単位を与えることを約束していた。この実験参加に対する異なる背景は，複数テキストの読解に影響した可能性がある。複数テキストを読解する際，大学の単位に関わるレポート作成用の読解と，個人的な事情のための読解では，前者のほうがテキスト間の読解切り替え数や，逆説的な接続詞の使用頻度などの統合的な理解を示す指標の数が多かった（Schoor et al., 2023）。つまり実験に真剣に取り組んだことにより，複数テキストの関係性をより正確に把握し，統合的な理解が深まった可能性がある。

第四に，テキストのカウンターバランスは部分的に行われていた。順序効果が完全には相殺できていなかったため，今回の結果に影響している可能性がある。完全なカウンターバランスのためには，テキストの条件×グループ数が必要である（Suzumi & Koizumi, 2021）。これにより1つ目のテストが次のテストに影響してしまう練習効果をおさえることが期待できる。

第五に，分析1においてテキストの主張要因と根拠要因間でつながり形成数に差が見られなかったのは，実験協力者数が影響している可能性がある。Plonsky（2015）によると，実験協力者のサンプル数が増えるほど，有意差が出やすくなる。List et al.（2021a）の実験協力者は89名であった一方，本研究の実験協力者は32名であった。

6.3 教育的示唆

本研究の教育的示唆は次の通りである。第一に，タスク介入の必要性である。本研究ではEFL学習者の熟達度が高い場合，複数テキスト間のつながり形成数が多い一方，熟達度が低い場合は複数テキスト間のつながり形成数が少ないことが分かった。複数テキストの統合的な理解を得ることはL1読者にとっても挑戦的であり，L2学習者には往々にしてテキスト間のつながりを無視する傾向があることを踏まえると，低熟達度のEFL学習者に対し，複数テキストの読解時にテキスト間のつながり形成を支援するタスク介入を行うことが望ましいだろう。

L1研究において複数テキストのつながり形成を支援する読解教示やタスクが研究されているので，EFL学習者への指導に応用することも検討に値する。読解教示（reading instructions）とは，読み手に読みの目的を与えることである（e.g., Horiba, 2013）。例えば，読解教示では，L1研究においてテキスト間の関係性を特定させる処理を促す教示により，ライティングでの統合の質が向上した（Kobayashi, 2009）。またタスク介入では，読解中にテキストをハイライトするタスクは読解理解を助けるといわれる。Leroy et al.（2021）は複数テキスト読解時にテキストのハイライトをさせるタスクを課した。テキスト間での類似性や差異性を特定させ，それぞれを異なる色でハイライトのマーキングをさせた場合，ライティングで

のテキスト間のつながりが多く含まれ，読み手のテキストの統合的理解を促進させることを発見した。このようにL1研究では，複数テキスト読解時にテキストのつながりや関係性を特定させる教示やタスク介入を与えた場合，読み手の統合的理解が深まることがわかっている。L2学習者が複数テキストのつながり特定や関係性のハイライティングタスクを行った場合，統合が促進される可能性がある。

第二の教育的示唆は，新たな作問開発の可能性についてである。EFL学習者の熟達度と複数テキスト間のつながり形成数の関係に基づき，テスト作成者は複数テキスト理解を問う問題を作成する場合，テキスト間で形成しうるつながりの数や種類，引き出す問いを操作するなど工夫することにより，複数テキスト理解問題の難易度を調整することができるかもしれない。例えば弁別力の高い項目に，つながり形成の数が考えられる。本研究では高熟達度学習者は複数テキスト間のつながり形成数が多い一方，低熟達度学習者はつながり形成数が少なかった。よって複数テキスト間に形成されうるつながり数を増やすことによって，読解問題の難易度を高めることができる可能性がある。

謝辞

本研究を実施する機会を与えてくださった公益財団法人 日本英語検定協会の皆様と選考委員の方々に心より感謝申し上げます。特に研究助言者である筑波大学の小泉利恵先生には，本研究で得られた分析データの解釈や報告書作成について貴重なアドバイスを頂戴し，深く感謝申し上げます。また筑波大学の卯城祐司先生には，本研究の着想段階から研究実施，報告書の作成に至るまで親身なご指導を頂戴しました。最後に，本研究にご協力いただきました皆様に厚く御礼を申し上げます。

引用文献

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review, 30*(3), 973–999. https://doi.org/10.1007/s10648-018-9436-8

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beker, K., Jolles, D., Lorch, R. F., & Van den Broek, P. (2016). Learning from texts: Activation of information from previous texts during reading. *Reading and Writing, 29*(6), 1161–1178. https://doi.org/10.1007/s11145-016-9630-3

Bloom, B. S., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Pearson.

Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-Induced Source Comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist, 52*(3), 167–181. https://doi.org/10.1080/00461520.2017.1323219

Braasch, J. L. G., & Kessler, E. D. (2021). Working toward a theoretical model for source comprehension in everyday discourse. *Discourse Processes,58*(5), 449-467. https://doi.org/10.1080/0163853X.2021.1905393

Braasch, J. L., McCabe, R. M., & Daniel, F. (2016). Content integration across multiple documents reduces memory for sources. *Reading and Writing, 29*(8), 1571–1598. https://doi.org/10.1007/s11145-015-9609-5

Braasch, J. L. G., Rouet, J. F. G, Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition, 40*(3), 450–465. https://doi.org/10.3758/s13421-011-0160-6

Bråten, I., & Braasch, J. L. G. (2018). The role of conflict in multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 184–201). Routledge.

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The Role of epistemic beliefs in the comprehension of multiple expository texts: toward an integrated model. *Educational Psychologist, 46*(1), 48–70. https://doi.org/10.1080/00461520.2011.538647

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J. F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative, comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209–233). Erlbaum.

Britt, M. A., Rouet, J.-F., & Braasch, J. L. G. (2013). Documents as entities: Extending the situation model theory of comprehension. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading—from words to multiple texts* (pp. 160–179). Routledge.

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading, 42*(2), 288–325. https://doi.org/10.1111/1467-9817.12269

Delgado, P., & Salmerón, L. (2022). Cognitive effort in text processing and reading comprehension in print and on tablet: an eye-tracking study. *Discourse Processes, 59*(4), 237–274. https://doi.org/10.1080/0163853X.2022.2030157

引用文献

Ferguson, L. E., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2013). Epistemic beliefs and comprehension in the context of reading multiple documents: Examining the role of conflict. *International Journal of Educational Research, 62,* 100–114. https://doi.org/10.1016/j.ijer.2013.07.001

平井明代・岡秀亮・草薙邦広(2022).『教育・心理系研究のための R によるデータ分析: 論文作成への理論と実践集』 東京書籍

Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language, 25,* 98–125. https://scholarspace.manoa.hawaii.edu/items/573df866-e067-45ab-9a97-450f4054dee3

Johns, A. M., & Mayes, P. (1990). An analysis of summary protocols of university ESL students. *Applied Linguistics, 11,* 253–271. https://doi.org/10.1093/applin/11.3.253.

Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing,* 25, 4–22. https://doi.org/10.1016/j.jslw.2014.05.005

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge University Press.

Kobayashi, K. (2009). The influence of topic knowledge, external strategy use, and college experience on students' comprehension of controversial texts. Learning and *Individual Differences, 19*(1), 130–134. https://doi.org/10.1016/j.lindif.2008.06.001

Lenth, R (2023). *emmeans: Estimated marginal means, aka least-squares means.* R package (Version 1.8.5). https://CRAN.R-project.org/package=emmeans.

Leroy, C., Gerjets, P., Oestermeier, U., & Kammerer, Y. (2021). Simultaneous presentation of multiple documents and text-highlighting: Online integrative processes and offline integrated understanding. *Scientific Studies of Reading, 25*(2), 179–192. https://doi.org/10.1080/10888438.2020.1784903

List, A. (2020). Investigating the cognitive affective engagement model of learning from multiple texts: A structural equation modeling approach. *Reading Research Quarterly, 56*(4), 781-817. https://doi.org/10.1002/rrq.361

List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist, 54*(1), 20–39. https://doi.org/10.1080/00461520.2018.1505514

List, A., Du, H., & Lee, H. Y. (2021a). Examining relation formation across consistent and conflicting Texts. *Discourse Processes, 58*(2), 134–154. https://doi.org/10.1080/0163853X.2020.1834328

List, A., Du, H., & Lee, H. Y. (2021b). How do students integrate multiple texts? An investigation of top-down processing. *European Journal of Psychology of Education, 36*(3), 599–626. https://doi.org/10.1007/s10212-020-00497-y

List, A., Du, H., Wang, Y., & Lee, H. Y. (2019). Toward a typology of integration: Examining the documents model framework. *Contemporary Educational Psychology, 58,* 228–242. https://doi.org/10.1016/j.cedpsych.2019.03.003

Lüdecke, D. R. (2023). sjPlot: Data visualization for statistics in social science. R package (Version 2.8.13). https://CRAN.R-project.org/package=sjPlot

三上洋介 (2022). タスク要因とテキスト要因が複数テキスト読解の統合的理解に与える影響, 全国英語教育学会, 第47回全国英語教育学会北海道研究大会予稿集, 240-241, 2022年8月7日

文部科学省 (2018). 高等学校学習指導要領(平成30年告示)解説 総則編. 東洋館出版社

日本英語検定協会(n.d.). 英検と CEFR との関連性について 研究プロジェクト報告. https://www.eiken.or.jp/eiken/group/result/pdf/report_02.pdf

日本学術会議 文学委員会 文化の邂逅と言語分科会 (2016). 提言:言語・ことばに対する能動的態度を育てる取り組み―初等中等教育における英語教育の発展のために― https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-t236.pdf

Perfetti, C. A., Rouet, J. F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. Van Oostendorp, & S. R. Goldman (Eds.), *The construction of mental representation during reading* (pp. 99–122). Erlbaum

Plakans, L. (2009). Discourse synthesis in integrated second language writing assessment. *Language Testing, 26*(4), 561–587. https://doi.org/10.1177/0265532209340192

Plonsky, L. (2015). *Advancing quantitative methods in second language research.* Routledge.

Primor, L., Yeari, M., & Katzir, T. (2021). Choosing the right question: The effect of different question types on multiple text integration. *Reading and Writing, 34*(6), 1539–1567. https://doi.org/10.1007/s11145-021-10127-8

Qin, J., & Liu, Y. (2021). The influence of reading texts on L2 reading-to-write argumentative writing. *Frontiers in Psychology, 12,* 655601. https://doi.org/10.3389/fpsyg.2021.655601

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rouet, J.-F., Le Bigot, L., de Pereyra, G., & Britt, M. A. (2016). Whose story is this? Discrepancy triggers readers' attention to source information in short narratives. *Reading and Writing, 29*(8), 1549–1570. https://doi.org/10.1007/s11145-016-9625-0

Rouet, J.-F., Saux, G., Ros, C., Stadtler, M., Vibert, N., & Britt, M. A. (2021). Inside document models: Role of source attributes in readers' integration of multiple text contents. *Discourse Processes, 58*(1), 60–79. https://doi.org/10.1080/0163853X.2020.1750246

Schedl, M., O'Reilly, T., Grabe,W., & Schoonen, R. (2021). Assessing academic reading. In X. Xi & J. M. Norris (Eds.), *Assessing academic English for higher education admissions* (pp. 22–60). Routledge. https://doi.org/10.4324/9781351142403-2

引用文献

Schoor, C., Rouet, J.-F., & Britt, M. A. (2023). Effects of context and discrepancy when reading multiple documents. *Reading and Writing, 36*(5), 1111–1143. https://doi.org/10.1007/s11145-022-10321-2

Saux, G., Britt, M. A., Vibert, N., & Rouet, J. (2021). Building mental models from multiple texts: How readers construct coherence from inconsistent sources. *Language and Linguistics Compass, 15*(3), 1-19. https://doi.org/10.1111/lnc3.12409

Stadtler, M., Scharrer, L., & Bromme, R. (2020). How relevance affects understanding of conflicts between multiple documents: An eye - tracking study. *Reading Research Quarterly, 55*(4), 625–641. https://doi.org/10.1002/rrq.282

Stang Lund, E., Bråten, I., Brandmo, C., Brante, E. W., & Strømsø, H. I. (2019). Direct and indirect effects of textual and individual factors on source-content integration when reading about a socio-scientific issue. *Reading and Writing, 32*(2), 335–356. https://doi.org/10.1007/s11145-018-9868-z

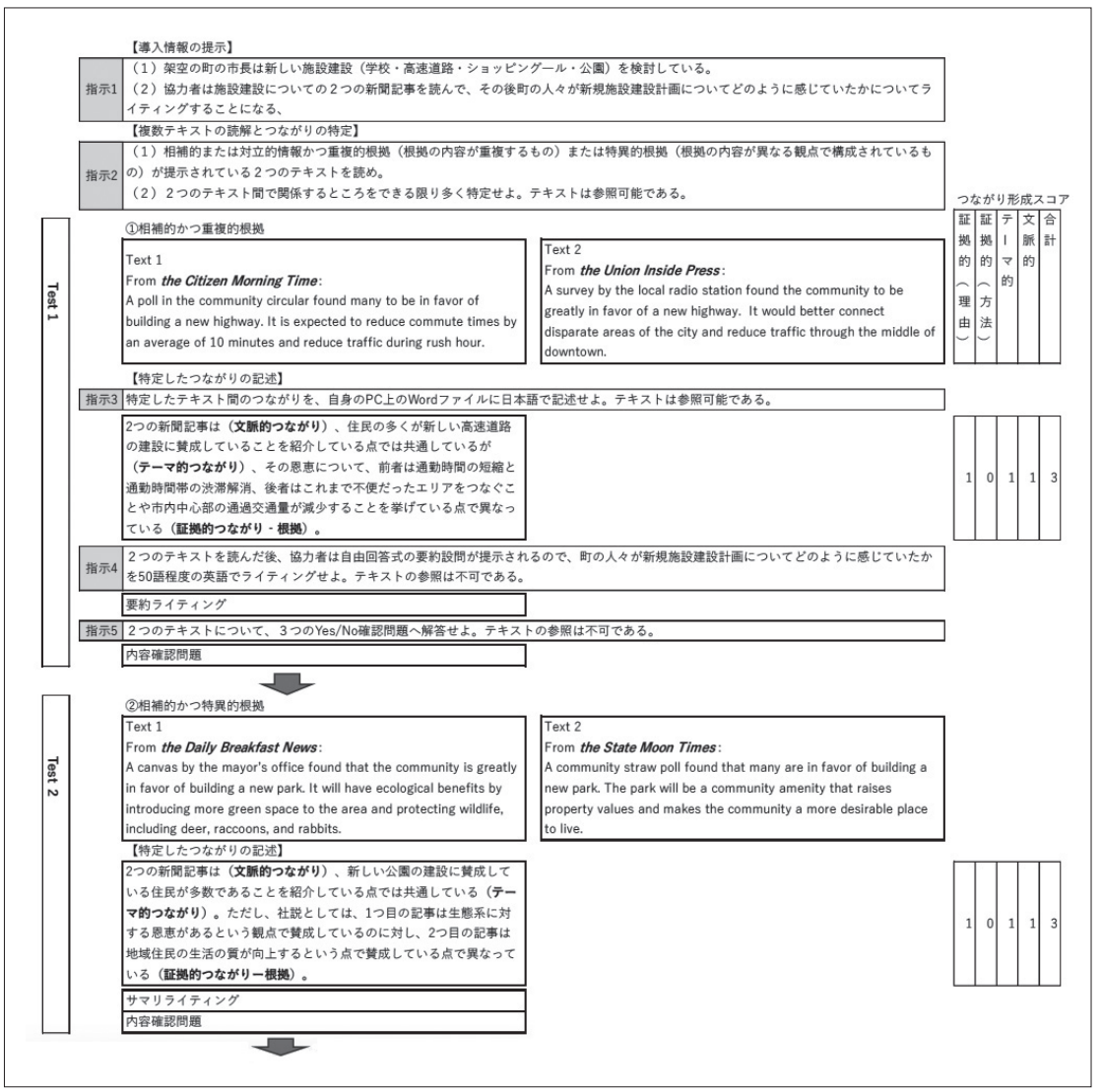
Suzuki, Y., & Koizumi, R. (2021). Using equivalent test forms in SLA pretest-posttest design research. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 457–467). Routledge.

van den Broek, P., & Kendeou, P. (2015). Building coherence in web-based and other non-traditional reading environments: Cognitive opportunities and challenges. In R. J. Spiro, M. DeSchryver, M. S. Hagerman, P. M. Morsink, & P. Thompson (Eds.), *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (pp. 104–114). Routledge.

van den Broek, P., Risden, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch Jr. & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353–373). Erlbaum.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*(1),73–87. https://doi.org/10.1037/0022 - 0663.83.1.73

資料1：つながり形成スコアの例



資料1：つながり形成スコアの例

Test 3

③対立的かつ重複的根拠

Text 1	Text 2
From the Evening Star Press: At a recent town hall, the community was greatly in favor of building a new school. The new school is expected to increase property values by 10%, making your town a more desirable place to live.	From the Morning Sun Times: A recent community survey found that many opposed building a new school. Building the school will increase taxes. It's projected that each household will have to pay additional $1,000 each year to fund the new building.

【特定したつながりの記述】

2つの新聞記事は（**文脈的つながり**）、新しい学校の建設に対する地域住民の評価を示しているが、前者は大半が賛成、後者は大半が反対であると言っており、内容が異なっている（**テーマ的つながり**）。しかし、学校の建設により地価が上昇することを示唆している点では共通している。

0	0	1	1	2

サマリライティング

内容確認問題

Test 4

④対立的かつ特異的根拠

Text 1	Text 2
From the Business Daily Press: A community meeting found the community to be greatly in favor of a new shopping mall. It will introduce new stores and more shopping options to the area, making it a more desirable place to live.	From the Local Voice Times: In interviews with the local paper, many were found to be opposed to the shopping mall. It will give the town a more commercial, rather than charming feel, and construction is expected to take a long time, increasing noise pollution.

【特定したつながりの記述】

2つの新聞記事は（**文脈的つながり**）、施設建設計画に対する住民の意見を記述しているところが似ている。前者は賛成派の立場である対し、後者は反対派の立場である（**テーマ的つながり**）。また、前者は賛成派のほうが、後者は反対派のほうが圧倒的に多いかのように表現している。

0	0	1	1	2

サマリライティング

内容確認問題

テキスト条件別スコア

①相補的かつ重複的根拠	1	0	1	1	3
②相補的かつ特異的根拠	1	0	1	1	3
③対立的かつ重複的根拠	0	0	1	1	2
④対立的かつ特異的根拠	0	0	1	1	2

主張要因・根拠要因別の合計スコア

相補的主張　①+②	2	0	2	2	6
対立的主張　③+④	0	0	2	2	4
重複的根拠　①+③	1	0	2	2	5
特異的根拠　②+④	1	0	2	2	5

資料2：統合的要約ライティングの採点例

0 点（テキスト間のメインアイディアが正確に統合されていない）

Moring sun said building highway make savings time average 10 minutes from traffic jam during rush hour. Evening press said building highway make cost $ 3 million. It will draw funds from other community.

1 点（メインアイディアが正確に統合されている）

The people of my town think that they are in favor of structing new school. They think good that the new school will provide students good education.

2 点（メインアイディアとソースからのサポート情報が正確に統合されている）

People in the town have favourable attitude toward or opinions on building new facility. One thinks it is good thing as people can use more rooms to improve their learning environment and efficiency and the other thinks people can use facility for AP classes or other outside activities.

3 点（メインアイディアが正しく統合され , サポート理由も提供され , この情報を一致する新聞ソースと属性づけている）

According to the two newspaper many locals are in favor of building a new park. The papers had taken a survey to collect the information. The reason for why the locals were in favor of building a park varied. Some were in favor because they believed that the park would bring more wildlife back. Others believed that the park would make the value of the surrounding areas higher making the community a more desirable place to live.

B 実践部門 ● 報告Ⅰ ● 英語能力向上をめざす教育実践

高専生による説明文理解を深める協同的な再話活動

― 状況モデル構築につながる質問項目の比較･検討 ―

研究者:伊東　賢　茨城県／茨城工業高等専門学校　講師･茨城県／筑波大学大学院　在籍

《研究助言者:西垣 知佳子》

概要　本研究の目的は，英語による再話を授業に取り入れた場合に，(a)発話流暢性が向上するのか，(b) 英語による再話が内容理解をさらに深めるのか，(c)どのような足場がけの質問が再話者の状況モデル構築につながるのかを検証することである。国立高等専門学校の学生が5週にわたり，英語による再話を取り入れた授業に参加した。分析の結果，再話中のポーズ時間は減少し，発話量，発話スピードは一定程度向上するが，言いよどみ数は変化しないことが示された。また英語による再話では，学習者はテキストの表現そのものに注意を向けるようになり，テキストベースの記憶が強化され，テキストの内容にそぐわない情報に気づく可能性が高まることが示された。一方で，状況モデルの構築を促す効果は確認できなかった。再話が滞った際の足場がけの質問については，授業者が用意した質問よりも，聞き手自身が語彙レベルで記憶の想起を促す質問を考えて行うことが多かった。足場がけの質問はテキスト情報をより多く再生することに寄与するが，必ずしも滞った再話を再開させることが，再話者がより豊かな状況モデルを構築することにつながるわけではないことが示された。

1 はじめに

再話とは，ある文章を読んだ後に，その文章を見ない状態でその内容を知らない人に語る活動である(卯城他，2009)。高等学校における英語の授業において，スピーキング，リスニング活動の不足が指摘され，4技能統合型授業の必要性が指摘される中，授業改善の一つの方法として，再話が注目されている(佐々木，2020)。

日本人英語学習者にとって，再話を授業に取り入れる目的と，英語と日本語のどちらで再話を行うのかという問題は密接に関係している(鷲見，2015)。再話がもたらす効果の一つとして，発話流暢性の向上が挙げられるが(佐々木，2020)，この場合，英語で再話を行うことが前提となる。一方で，Kai(2008, 2009)は，日本語で再話を行うことで，再話を行わない場合よりも読解した内容の理解が深まることを示した。学習した内容の理解を深めるために再話を行う場合，再話は日本語で行われることが多い。英語で再話を行う場合，学習者の英語熟達度によっては，極端に発話量が制限され，再話そのものが成立しない可能性があるからである。

英語力の向上を目的とした場合，英語で再話を行うことに異論の入る余地はないであろう。一方で，日本語による再話と同様，英語による再話が内容理解に何らかの影響を及ぼす可能性はあるのだろうか。本研究では，英語による再話を授業に取り入れた場合，(a)発話流暢性は向上するのか，(b)読解したテキストの内容理解に影響を及ぼすのか，(c)再話が滞った際に聞き手が行う足場がけとしての質問は，内容理解に影響を及ぼすのかについて検証する。

2 先行研究

2.1 読み手が構築する心的表象

テキストを読解する際に，読み手はその内容について心的表象（mental representation）を構築する（W. Kintsch, 1988; Zwaan & Radvansky, 1998）。W. Kintsch（1988）が提唱する構成・統合モデル（Construction-Integration model）によると，読み手が構築する心的表象には以下の3つのレベルが存在する（W. Kintsch, 1988, 2013; W. Kintsch & E. Kintsch, 2005）。その一つは，表層的記憶（the surface-level memory）と呼ばれるもので，テキストの記述内容を，表記通りに記憶した表象に相当する。次に，命題的テキストベース（the propositional textbase; 以下，テキストベース）は，命題レベルでテキストに述べられた内容を理解し記憶することに相当する。命題とは，その内容の真偽を判断することができる知識の最小単位と定義される（卯城他，2009）。つまり，テキストベースレベルの心的表象とは，テキストの表現そのものを記憶しているのではなく，そのテキストが伝えようとする内容や意味に焦点を当てて，命題単位の理解に基づいて構築される表象に相当する。最後に，状況モデル（the situation model）は，読み手が推論を行い，テキストに書かれている情報と読み手の背景知識とを統合して達成される深い理解に基づいて構築される表象である。最終的な読解の目的は，読解した内容について首尾一貫した状況モデルを構築することであり（Zwaan & Radvansky, 1998），そのためには，テキスト内で明示的に示されている情報を理解するだけではなく，背景知識をもとに推論を行い，明示的に示されていない情報や情報間のつながりについても理解することが必要となる（W. Kintsch, 1994; W. Kintsch & E. Kintsch, 2005; McNamara, E. Kintsch, Songer, & W. Kintsch, 1996）。

2.2 再話の効果

英語教育研究における再話がもたらす効果については，数多く報告されている。例えば，再話を行うことで，英語の語彙や統語規則の習得につながり（e.g., Joe, 1998），英語で発話する際の流暢性（fluency）が向上する（e.g., 佐々木，2020）。これらの効果を享受するためには，英語で再話を行うことが前提となる。学習者の英語の使用量を増やすことが，英語運用に関わる知識の習得や流暢性の向上につながると考えられる。

また，再話を行うことにより，内容理解が深まることが報告されている（e.g., Gambrell, Koskinen, & Kapinus, 1991; Kai, 2008, 2009; Wilson, Gambrell, & Pfeiffer, 1985）。Kai（2008）は，物語文の読解後に再話を行う群と，再話を行わない群とを比較し，再話を行う群のほうが，内容理解が深まることを示した。また，Kai（2009）は，オリジナルの文章，主題文を除いた文章，主題文の位置を移動させた文章を用意し，再話の効果を検証した。検証の結果，主題文の操作により主題の把握が難しくなった文章において，再話が大局的一貫性（global coherence）の構築を促す可能性が示唆された。これらの内容理解に及ぼす再話の効果を享受するためには，基本的に再話は日本語で行われる。これは，日本人英語学習者が英語で再話を行う場合，英語の運用能力の制約により，日本語による再話ほど流暢に行うことが難しいことに起因している。

以下，本研究で取り扱う（a）再話が流暢性向上につながる理由と，（b）再話が内容理解を深める理由について，その理論的背景を紹介する。

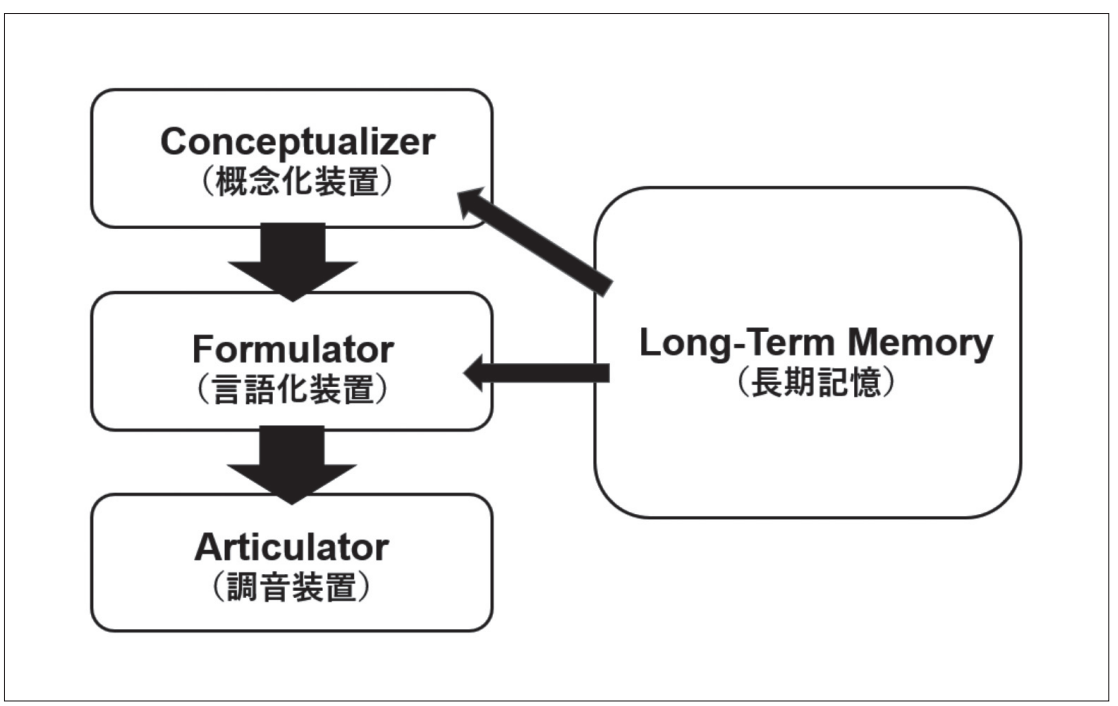
2.2.1 再話が流暢性向上につながる理論的背景

流暢性の定義は一様ではないが，第二言語研究における流暢性は，即時的な処理が求められる状況下で，言語知識を容易に動員できることによって生まれる，発話の表面的な滑らかさと定義される（Foster,

2020)。流暢性は，発話流暢性(utterance fluency)，認知流暢性(cognitive fluency)，知覚流暢性(perceived fluency)の3つの構成概念から成り立っている(Segalowitz, 2016)。発話流暢性は，伝えるべき情報が，言いよどみやポーズがなくスムーズに発話されているかどうかに関する流暢性である。次に，認知流暢性は，発話に必要な認知的処理がスムーズに行われているかどうかに関する流暢性である。最後に，知覚流暢性は，実際の発話に対する主観的な評価に基づく流暢性である。発話流暢性は，発話スピードに関する speed fluency，ポーズの有無に関する breakdown fluency，言いよどみや繰り返しに関する repair fluencyの3つの領域に分類される(Housen, Kuiken, & Vedder, 2012)。

Kormos(2006)が提唱する第二言語学習者の発話プロセスに関するモデルによると(図1)，発話に至るまでのプロセスには，(a)伝えたい内容をまとめる概念化(conceptualization)，(b)その内容を発話で用いる言語体系に則って言語化(encoding)，そして(c)言語化した内容の音声化(articulation)という3つの段階がある。これらはそれぞれ図1の Conceptualizer, Formulator, Articulator で行われるプロセスに相当する。第二言語における流暢性向上に向けて重要となるのは，情報を符号化するプロセスの自動化と慣用表現(formulaic language)の使用である(Kormos, 2006)。英語でスピーキング活動を継続的に行うことで，言語化のプロセスにおいて，宣言的知識が手続き的知識に変わり(Skehan, 1998)，結果として，プロセスの自動化が進み，流暢性の向上が期待できる(Kormos, 2006)。

再話では，発話内容は読解したテキストの内容に基づいており，テキストに登場した表現を使って再話することが多い。すなわち再話は，ゼロから発話内容を組み立てなくてはならない活動よりも，概念化，言語化に割く認知資源は少なくて済むと考えられる。結果として，その他の発話プロセスに余剰の認知資源を割り当てることができ，発話流暢性の向上につながる可能性がある。同時に，再話ではテキストの表現が多用されることになるため，慣用表現の習得にもつながると考えられる。



■図1：第二言語発話生成モデル(Kormos, 2006を参考に著者が作成)

2.2.2 再話が内容理解を深める理論的背景

再話が内容理解を深める理論的背景として，以下の3つの可能性を紹介する。一つは，再話を行うことで，学習者がテキストにより深く関与することにつながる可能性である。再話は，テキスト構造の理解を促し

(Gambrell et al., 1991; Koskinen, Gambrell, Kapinus, & Heathington, 1988; Morrow, 1985), より多くのテキスト情報の記憶につながる(Koskinen et al., 1988)。また, テキストの内容を全体的に再構築する手助けとなり(Koskinen et al., 1988), その内容を自身の経験と結びつけて理解することにつながる(Gambrell et al., 1991; Koskinen et al., 1988)。再話は学習者をテキスト構造に注目させ, テキスト内の各情報がどのように結びついて首尾一貫した内容を構成しているのかについて, より深い理解を促す可能性がある。

次に, 再話を行うことで, テスト効果(the testing effect)が現れる可能性である(Ito, in press)。Roediger, Putnam, and Smith(2011)によると, テストを受ける際に記憶から情報を取り出すという行為が, その内容の記憶を強化する。再話とは, 学習者が読解した内容を記憶し, テキストを見ない状態でその内容を産出する活動である。つまり再話を行うことで, 発話された内容に関する記憶がテスト効果により強化され, その結果, 内容理解が深まる可能性がある。この場合の内容理解とは, テキスト情報の記憶に相当すると考えられるため, テスト効果により, 再話がテキストベースレベルの理解に影響を及ぼす可能性がある。

最後に, 再話が読解した内容について推論情報の付加を促す可能性である(Ito, in press)。伊藤・垣花(2009)では, 聞き手に学習内容を説明することで(注[1]), 話者自身が学習した内容を修復し, その内容を再構成することが示され, これを「意味付与的説明」の効果とした(注[2])。「意味付与的説明」とは, テキストに記述された内容に, 発話者が意味や解釈を加えたものである。首尾一貫した状況モデルの構築には, 発話者が推論を働かせて, 自らの既有知識をテキスト情報と統合することが必要となるが(W. Kintsch, 1988; W. Kintsch & E. Kintsch, 2005), 「意味付与的説明」が再話にも現れるとすれば, 再話が状況モデルレベルの理解に影響を与える可能性がある。

2.3 協同的な再話活動と足場がけの質問

伊藤・垣花(2009)は, 聞き手の存在により, 説明する側の内容理解が促進される可能性を示したが, これは, 聞き手の存在を前提とする再話にも当てはまると思われる。一方で, 日本人英語学習者にとって, 英語による再話の場合, 伝えたい内容を十分に伝えることができない状況が起こりうる。再話が滞った際には, 再話を再開させるために促しの質問(prompting questions)を与えて支援することが効果的である(Tachibana, 2016; 小河原・木谷, 2020)。具体的に, 聞き手が質問やヒント, 発話の促しや修正を与えることで, 再話者の発話を支援することができる(小河原・木谷, 2020)。また, 聞き手がテキストのキーワードについて質問することが, 話す内容に気づかせる足場がけ(scaffolding)となる(Tachibana, 2016)。このような促しの質問による支援は, 特に英語で再話を行う場合に有効である。

再話を促す支援の効果は, 質問タイプによって異なる可能性がある。質問タイプには大きく分けて, テキストに明示的に示された情報を読み取って解答する事実質問(パラフレーズ質問)と, テキストには直接示されていない内容を, テキスト情報と背景知識から推測して解答する推論質問がある(清水, 2005; 田中・辻, 2015)。清水(2005)は, 英検, TOEFL, 大学入試センター試験の問題を分析し, 推論質問の方が事実質問よりも難易度が高いことを示した(注[3])。事実質問には, 質問そのものに, 解答に必要なキーワードが含まれていることが多く, 再話者は, 疑問詞に置き換えられた情報を想起するだけで質問に解答できることが多い。結果として, 情報を想起するための認知的負荷は小さいと考えられる。一方で, 推論質問は, 解答に必要な情報について, 推論を働かせて, 明示的に示されていない情報間のつながりも理解しなければならないため, より深い処理が必要となる。つまり, 事実質問と推論質問を足場がけの質問として使用する場合, 前者の方が, 再話者が想起すべき情報は少なく解答が容易で, 後者の方が, 再話者が処理しなければならないプロセスが多く, 解答が難しいと考えられる。van den Broek and Helder(2017)は, 読解後の活動が, 読解時に構築した心的表象をさらに変化させる可能性を指摘しているが, 再話者が滞った再話を再開しようとする際に, 質問タイプがもたらす情報処理の違いが, 再話者の心的表象に異なる影響を及ぼす可能性がある。

2.4 文再認課題

学習者の内容理解を測定する手法の一つに文再認課題がある。文再認課題とは，文章を読解後，その文章に登場した文と，一定の基準に基づき改編した文を提示し，それらの文が読解した文章に登場したかどうかを判定する課題である（e.g., Karimi & Richter, 2023; Schmalhofer & Glavanov, 1986; Wolfe & Woodwyk, 2010）。文再認課題を用いると，学習者の理解レベル（テキストベース，状況モデル）を同時に測定することができる（Schmalhofer & Glavanov, 1986）。学習者が構築した心的表象をテキストベースと状況モデルの各レベルで測定する場合，文再認課題は3種類のテスト項目から構成される（e.g., Karimi & Richter, 2023）。テスト項目の一つは「言い換え文（paraphrase）」と呼ばれるもので，同義語によるテキスト表現の言い換えや，能動態から受動態への変換といった操作によって作成される。つまり，本来の文と一字一句同一の表現ではないが，テキストベースレベルの情報は保持しているとみなされる文である。もう一つは「推論文（inference）」と呼ばれ，テキストに明示的に述べられていない情報を含んでいるが，適切な推論に基づいて理解すれば，テキストが伝えようとする内容と一致すると判断できる文である。つまり推論文は，テキストベースレベルではテキスト情報を保持していないが，状況モデルレベルではテキスト情報に一致するとみなされる文である。最後は「錯乱文（distractor）」で，テキストに登場した表現をもとに作成されているが，テキストの内容に矛盾する文である。

それぞれのテスト項目に対する再認確信度（各テスト項目が登場したか，しなかったのかを判定する際の確信の度合い）を測定することで，学習者が構築した心的表象の強度を，テキストベースと状況モデルのレベルごとに測定することができる。具体的に，言い換え文の再認確信度を測定することで，テキストベースの強度を測定することができ，推論文の再認確信度を測定することで，状況モデルの強度を測定することができる（Ito, in press; 邑本，2000）。

2.5 本研究の目的と研究課題

本研究の研究課題を設定する前に，先行研究の限界点について指摘したい。第一に，再話が内容理解を促進する効果が報告されているが，日本語による再話と英語による再話で，テキストベースと状況モデルに及ぼす影響が異なるのかどうかについては検証が進んでいない。これは，内容理解を深める効果を求める場合，多くの場合，再話は日本語で行われてきたからである。再話で使用する言語（日本語 vs. 英語）が異なると，内容理解に与える影響は異なる可能性がある。さらに，協同的な再話活動において，支援の一つとして足場がけの質問を与える場合，どのような質問が支援として効果的なのかについても検証が必要である。

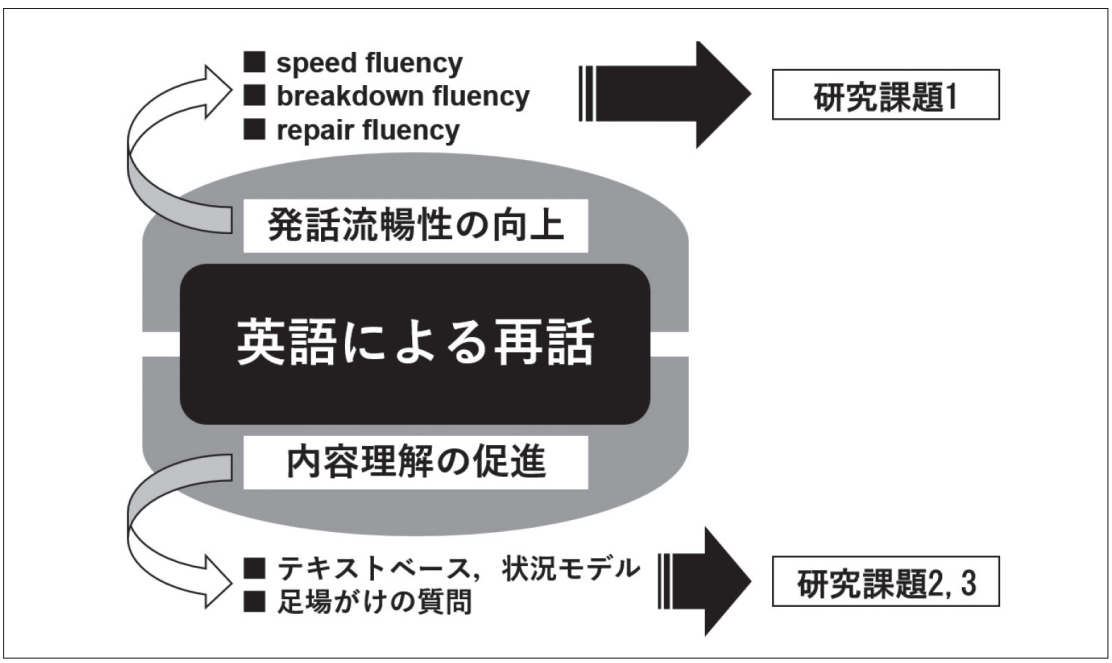
よって本研究では，初めに，英語による再話を行うと発話流暢性が向上し，内容理解が促進されるのかについて検証する。発話流暢性については，speed fluency, breakdown fluency, repair fluency の観点から分析する。内容理解に及ぼす効果については，W. Kintsch（1988）に従い，テキストベースレベルと状況モデルレベルごとに，理解への影響を検証する。また，足場がけの質問については，質問の質を統制して，どのような質問が状況モデルの構築を促す再話につながるのかについて検証する。よって，本研究では，以下の3つの研究課題を設定する。

研究課題1　英語による再話を取り入れると，発話流暢性は向上するのか。

研究課題2　英語による再話は，テキストベースレベル，状況モデルレベルの内容理解に影響を及ぼすのか。

研究課題3　再話が滞った際に足場がけの質問を与えると，テキストベースレベル，状況モデルレベルの内容理解は深まるのか。その場合，どのような質問がより効果的なのか。

本研究課題に関わる要因は，図2に示されている通りである。研究課題1については，英語による再話を取り入れることで，学習者は英語で発話をする機会が増え，その結果，発話したい内容を英語に符号化する段階で自動化が促進され，発話流暢性は向上すると考えられる。研究課題2については，英語による再話は，日本語による再話よりも発話量が制限される可能性が高いが，事前に十分な準備を行って再話に取り組めば，内容理解を深める効果が現れる可能性がある。研究課題3については，足場がけの質問に答える過程を経ることで，内容理解を深める効果が現れると考えられる。この点については，足場がけの質問の効果が確認されたうえで，質問の質がもたらす違いについても検証する。



■図2：本研究に関わる要因及び研究課題との関係

3 調査方法

3.1 協力者

協力者は国立高等専門学校に通う日本語を母語とする学生である。協力者は2年生に在籍し，これは高等学校2年生に相当する。協力者の専攻は，機械・制御，電気・電子，情報，化学・生物・環境に関わる4つの分野である。本研究は5週にわたる調査になるため，調査開始時の協力者数は194名であったが，最終的な協力者数は140名であった（注[4]）。事後アンケートの結果，協力者の英語熟達度は CEFR レベルで A1 から A2 レベルであると推定された（英検2級：n = 5，英検準2級：n = 17，英検3級：n = 80，取得級なし：n = 38）。

3.2 マテリアル

3.2.1 読解熟達度テスト

本研究は，協力者間計画で実施されるため，条件間で協力者の英語熟達度を統制するために，読解熟達度テストを用意した。具体的に，英検2級，準2級の過去問を使用して読解熟達度テストを作成した（旺文社，2010a，2010b）。1問につき1点の配点とし，26点満点のテストとした。

3.2.2 読解テキスト

読解テキストとして Breaking News English というウェブサイトから5つのテキスト（略称: FASHION, ROCKET, HATS, AUSTRALIA, PIGS）を選択した（Banville, 2021, 2022）。このウェブサイトの特徴として，各テキストにつき難易度別に複数のバージョンが用意されており，学習者の熟達度に応じたテキストを選択できる点が挙げられる。本研究では，学習者の熟達度を考慮して，選択した5つの文章は全てレベル0のものを使用した。

著者ともう1名の評価者（英語教育学専攻の博士課程の学生）が，Ikeno（1996）に従い，テキストをアイデアユニット（IU）に分割した。2名の評価者が独立して分割を行い，評価者間一致率を測定した（FASHION; 97.00%, ROCKET; 95.00%, HATS; 96.36%, AUSTRALIA; 100%, PIGS; 95.28%）。IU を分割する箇所について疑問が生じた場合には，すべて話し合いを経て解決した。各テキストの単語数，IU 数，読みやすさを示す指標の一つである Flesch-Kincaid Grade Level（FKGL）は表1の通りである。

■表1: 本研究で使用したテキスト

タイトル	略称	使用週	単語数	IU数	FKGL
Fast fashion is harming the environment	FASHION	第1週	101	19	5.3
Bigger risk of being killed by falling space rocket	ROCKET	第2週	122	20	4.0
Anti-cheating exam hats used in Philippines college	HATS	第3週	112	21	4.8
Australia's 'shocking' decline in environment	AUSTRALIA	第4週	111	20	6.1
Pigs are intelligent enough to play video games	PIGS	第5週	106	23	5.5

研究課題1では，第1週目と第5週目の再話データを比較するため，この2回の授業では，できるだけ難易度が同レベルのテキストを選択した。学習者の背景知識に応じて読みやすさは一様ではないが，FKGL から判断すると，FASHION と PIGS の読みやすさがおおよそ同レベルと考えられるため，FASHION を第1週の教材，PIGS を第5週の教材とした。残りのテキストの使用順序については，FKGL から判断した読みやすさが徐々に難しくなるように配列した。最終的に ROCKET を第2週，HATS を第3週，AUSTRALIA を第4週の教材とした。

3.2.3 文再認課題

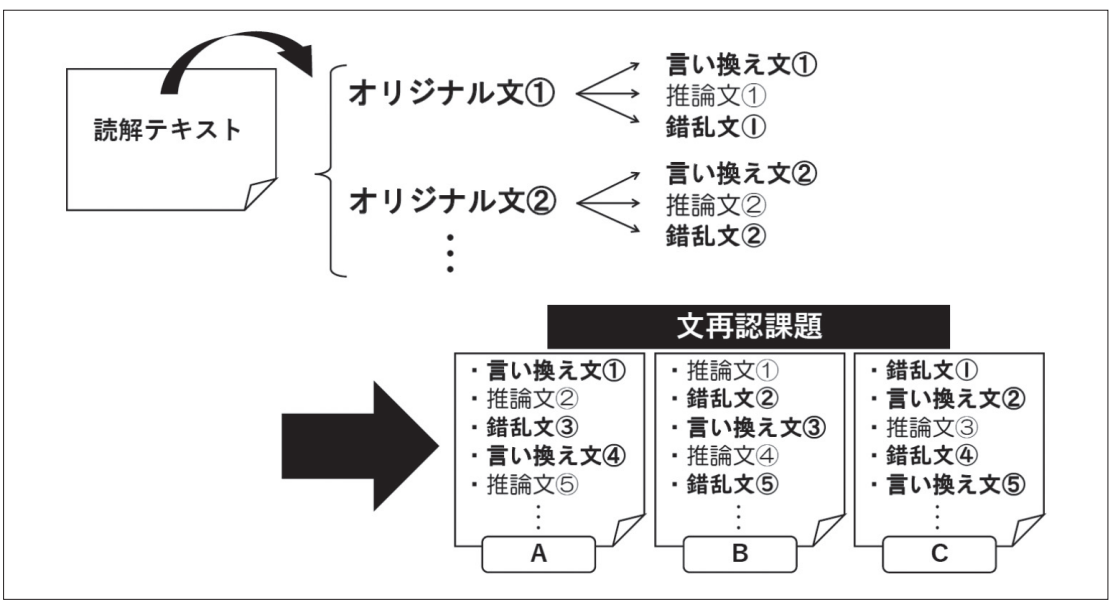
本研究では，協力者の内容理解を測定するために文再認課題を使用した。できるだけテキスト表現の記憶が再認判断に影響を及ぼさないようにするために，文再認課題は日本語で作成した（邑本，2000）。協力者のテキストベースと状況モデルのそれぞれのレベルの理解を測定するために，言い換え文，推論文，錯乱文の3つのテスト項目を用意した（e.g., Karimi & Richter, 2023; Wolfe & Woodwyk, 2010）。まず，テキストの一文一文（オリジナル文）が上記3つのテスト項目を作成する候補となった。言い換え文は，邑本（2000）に従い，オリジナル文を文字通りに日本語に訳した文とした。推論文と錯乱文は Karimi and Richter（2023）に基づいて作成した。具体的に，推論文は「もともとの文には明示されていないが，妥当な推論を行えば，その内容がテキストの内容と逸脱しないと判断できるもの」，錯乱文は「表層的にはテキストに使用されている用語を使用しているが，テキストの内容から逸脱したもの」とした。初めに，著者が全てのテスト項目を作成し，その後，別の評価者がテスト項目の質を評価した。すべてのテスト項目が作成基準に基づいて妥当かどうかを議論し，不一致箇所については議論を経て修正し，テスト項目を完成させた。具体例は，表2の通りである。

■表2: 文再認課題の言い換え文, 推論文, 錯乱文の例

テスト項目	例
言い換え文	学生たちは, 他の学生のテスト用紙を見ることができません。
推　論　文	学生たちは, 帽子をかぶっていると, 不正行為はできません。
錯　乱　文	学生たちは, 帽子をかぶっていても, 不正行為をしようとします。

注. オリジナル文は The students cannot look at other students' papers (Banville, 2022, October 27).

3種類のテスト項目(言い換え文・推論文・錯乱文)を作成したのちに, 以下の手順でテスト項目を配列し, 文再認課題を作成した。同一のオリジナル文に基づく3種類のテスト項目の一つだけが, 文再認課題に登場するように, 一つのテキストにつき3つの文再認課題(A～C)を作成した。これは同一のオリジナル文に基づくテスト項目同士を比較できないようにするためである。図3の手順に従って, 言い換え文, 推論文, 錯乱文の数は, 文再認課題(A～C)ごとにできるだけ等しくなるように配列した。各テキストについて, テスト項目作成の候補となるオリジナル文は12~13文であったため, 文再認課題も12~13個のテスト項目で構成された。最後に, テスト項目の提示順がランダムになるように並べ替えた。



■図3: 文再認課題の作成手順

3.2.4 足場がけの質問

研究課題2, 3では, 第2~4週のデータを使用するために, ROCKET, HATS, AUSTRALIAの3つのテキストについて, 事実質問と推論質問を用意した。事実質問は, 疑問詞で置き換えられた情報を文章から探して, その情報を疑問詞の部分に当てはめることができれば解答できる質問で, 基本的に what で始まる疑問文とした。一方推論質問は, その質問に対する理由を解答する質問で, 原因・結果の因果関係を推論して理解しなければ解答できない質問とした。具体的に推論質問は why で始まる疑問文とした。最初に著者が全ての質問を作成し, その後, 別の評価者が質問の質を評価した。すべての質問が作成基準に基づいて妥当かどうかを議論し, 不一致箇所については議論を経て修正し, 事実質問と推論質問を完成させた(資料1)。

具体例を表3に示す。それぞれの質問は A4用紙に印刷され，1回目の再話を行う前に協力者に配布された。

■表3: 事実質問・推論質問の例

質問タイプ	例
事実質問	Q. *What* must students in the Philippines wear in exams? A. They must wear *anti-cheating hats*.
推論質問	Q. *Why* must students in the Philippines wear hats in exams? A. *Because they cannot look at other students' papers*.

事実質問と推論質問を準備する一方で，協力者によっては，適切な場面でこれらの質問を適切に選択して再話者を支援することが難しいことも予想された。小河原・木谷（2020）は，質問の仕方や発話の促しは決して簡単ではないと報告している。よって，事実質問と推論質問の他に，日本語，英語を問わず，自由に足場がけの質問を考えて質問することを許可する条件も用意した。本研究ではこの質問を「自由質問」と名付けた。自由質問条件を設けることで，より多くの場面で質問による支援が確認できると考えた。自由質問は英語を基本とするが，質問する際の負担を減らすために，英語で質問することが難しい場合には日本語で質問することも可能とした。

3.3 手順

本研究開始前に，本研究の目的，方法，手順について協力者に説明した。特に，データの扱いは個人の特定ができないように配慮すること，データ分析の際には，承認を得られた学生のデータのみを利用すること，データの利用を承認しないからと言って，評価等について不利にならないことを周知した。

授業は，著者が担当する2年生5クラスで5週にわたって行われた。毎週50分の授業を2回実施し，合計100分で一つのテキストを学習した。そのうち4クラスについては，3時限目と4時限目のように2コマ連続であった。著者が5クラスすべての授業を担当した。5週にわたる指導実践の前に，読解熟達度テストを実施した。

個々の活動の実施手順については図4の通りである。前半50分では，初めにオーラルイントロダクションを行い，協力者はテキストを見ない状況で，授業者がパワーポイントを用いて，概要を英語で説明した。オーラルイントロダクションは再話に向けてモデルを示すという意味でも重要な活動である（語学教育研究所，2014）。次に，語彙の確認を行った。語彙の確認の後で，テキストや語彙，T/F クエスチョン等を載せたハンドアウトを配付し，協力者はリスニング，T/F クエスチョンに取り組んだ（注5）。T/F クエスチョンは，オーラルイントロダクションによる概要把握から，内容理解の上で重要な箇所に注意を向けることを意図して行われた。その後，プリントに従って詳細な内容理解に努めた。内容理解の際には，文章の構造に注目するように促した。また，後で行う再話の準備として，テキストの表現をどのように言い換えて表現できるかについても指導した。授業は基本的に英語で行い，協力者の様子を見ながら，適宜日本語で説明を付け加えた。

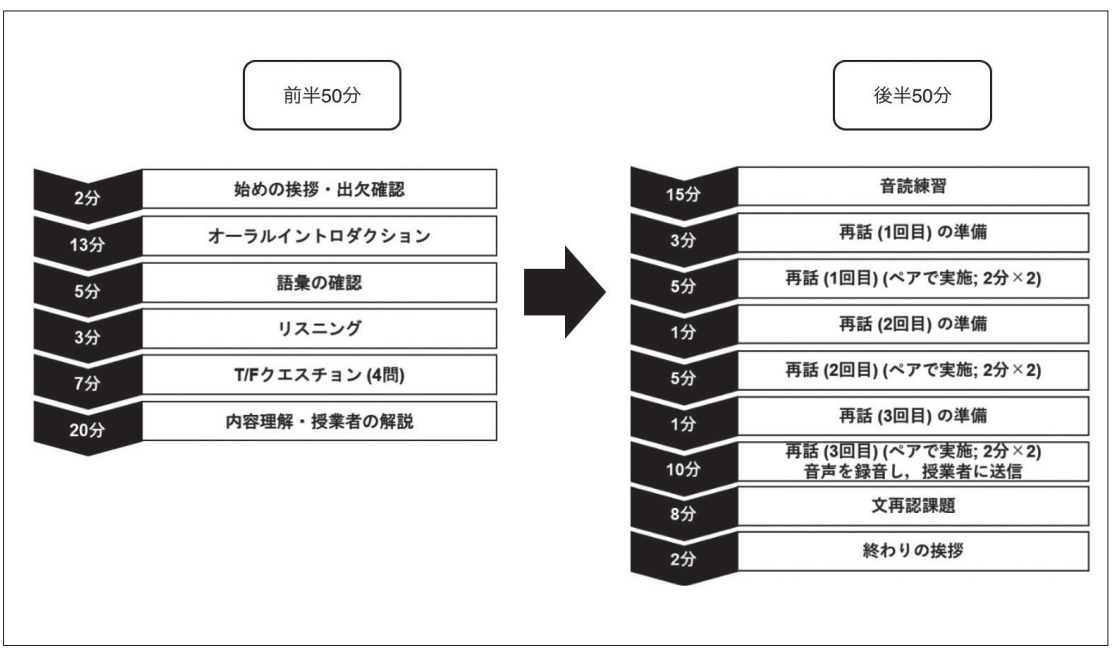
後半の50分では，音読練習，再話，文再認課題の順に実施した。音読練習では，授業者のモデルに合わせたモデルリーディングに続けて，バズリーディング，ペアリーディング，穴あきリーディング（スクリーンにいくつかの単語を黒塗りした文章を示し，黒塗りにされた単語を自分で補いながら行うリーディング），オーバーラッピングリーディング等で，十分に音読練習を行った。

音読練習後に，隣同士の学生でペアになって英語で再話を行った。再話はペアを変えて合計3回行った。再話する内容をまとめるために，第1回目の再話の前に3分間の準備時間を設けた。第2，3回目の再話の際

には，1分間の準備時間を設けた。授業で使用したハンドアウトにメモを取ることも許可したが，再話の際はテキスト，メモを見ることを禁止した。絵や図といった視覚情報も提供しなかった。再話の際には，テキストの表現をそのまま用いるのではなく，自分の言葉で伝えること，聞き手が理解できるように内容を整理して伝えること，できるだけ学習したことを多く伝えることを指示した。

合計3回の再話で，ペアのどちらが最初に始めるのかについては，授業者の指示に従い，再話を最初に始める場面が協力者間でできるだけ均等になるように配慮した。第3回目の再話を録音し，このデータが分析対象になった。録音は各協力者が所持するスマートフォンの録音機能を使用して行った。録音後に授業者のメールアドレス宛に録音データを送信させて，録音データを収集した。

録音データ送信後，文再認課題に取り組んだ。文再認課題は Google Forms を使用して実施した。一つのテキストから3種類の文再認課題を作成したが，協力者がどの課題に取り組むのかについては，授業者が前もって割り当てを決めておき，協力者に示した。これは，3種類の文再認課題に取り組む協力者数をおおよそ均一にするための処置である。協力者は各自のノートパソコンかスマートフォンを使用して，文再認課題に取り組んだ。文再認課題終了後，授業者が簡潔に授業のまとめを行い，授業は終了した。



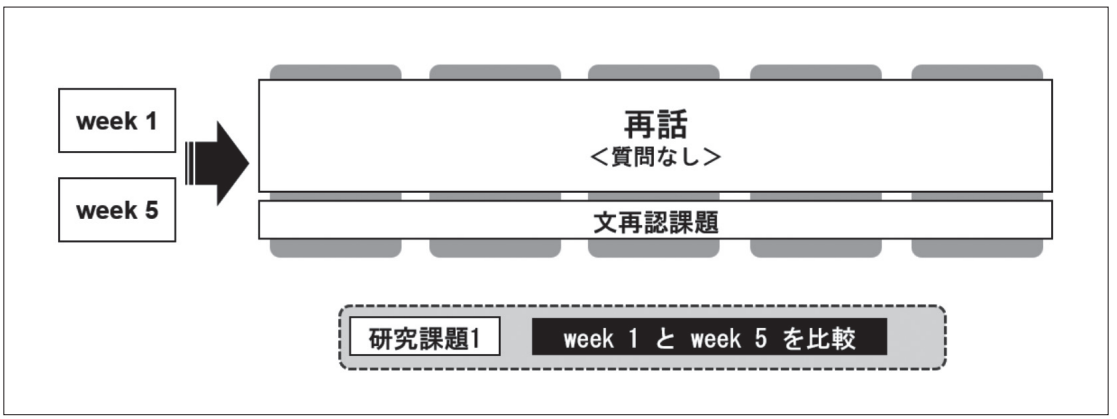
■図4：基本的な授業の流れ（左側が前半50分，右側が後半50分）。後述の通り，後半の指導手順については，第2週～第4週の期間は，再話時の質問条件及び，文再認課題と再話の実施順序が，クラスごとに異なる。

以上が基本的な授業の流れであるが，研究課題を検証するために，5クラスそれぞれに条件を設定し，その条件ごとに，活動の順番を入れ替えたり，足場がけの質問を与えたりといった変更を加えた。この点について以下に詳述する。

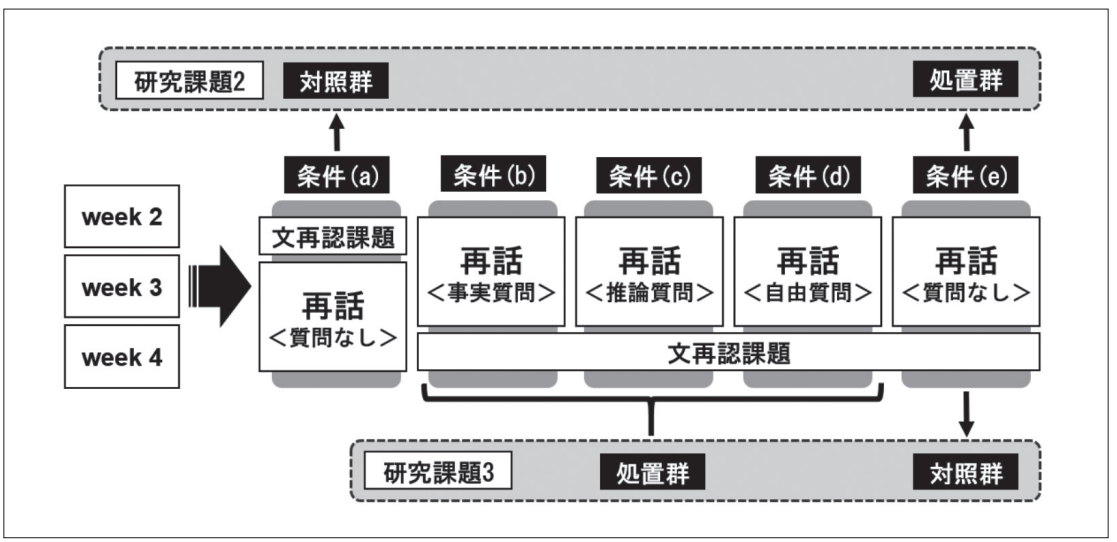
発話流暢性の検証（研究課題1）については，第1週と第5週のデータを比較するため，これらの週は，再話と文再認課題を行う手順は全クラスで統一した（図5）。具体的に，足場がけの質問を行わない条件で再話を行い，再話の後に文再認課題を実施した（注[6]）。

第2週～第4週については，英語による再話が内容理解に及ぼす影響（研究課題2）と，足場がけの質問が状況モデル構築に及ぼす影響（研究課題3）を検証するために，クラスごとに以下の5条件の一つを割り当てた（図6）。条件の一つは（a）再話を行う前に文再認課題を行うこととした。再話の前に文再認課題を行うことで，再話が文再認課題の結果に影響しないからである（注[7]）。残りの4条件として，（b）再話時に事

実質問を行う条件，(c)再話時に推論質問を行う条件，(d)再話時に自由質問を行う条件，(e)再話時に質問を行わない条件を設けた。条件(a)と条件(e)を比較することで，再話条件(再話あり・再話なし)による比較ができ(研究課題2)，条件(b)～条件(e)を比較することで，質問条件(事実質問・推論質問・自由質問・質問なし)の比較ができると考えた(研究課題3)。



■図5：第1週，第5週における再話条件と文再認課題の実施順序及び研究課題との関係



■図6：第２週～第４週における各条件及び研究課題との関係

3.4 分析

データは R(version 4.1.0, R Core Team, 2021)を用いて分析した。第1週と第5週の両方の授業に参加して，さらに，第2週から第4週の3回の授業において，最低1回は授業に参加した協力者を分析対象とした。結果として140名が分析対象になった。

授業実施前は，図5, 6の通り，各クラスに一つの条件を割り当ててクラス間の比較を行うことで，研究課題の検証を行う予定であった。しかし，条件(b)～条件(d)の足場がけの質問を行う条件で，再話が滞った際に，すべての聞き手が質問による支援を行ったわけではないことが明らかになった。そこで，一度でも聞き手が足場がけの質問を行い，その質問が新たなIUの再生につながった場合，滞っていた再話を再開させる支援になったと判断し，改めて協力者がいずれの条件に当てはまるのかを確認した(表4)。

■表4: 実験後に再編成した5つの条件ごとの協力者数

条件	対照群・処置群		授業実施前に割り当てた協力者数	実際に各条件に当てはまる協力者数
	研究課題2	研究課題3		
(a) 再話なし	対照群		39	29
(b) 再話あり(事実質問)		処置群	37	7
(c) 再話あり(推論質問)		処置群	41	2
(d) 再話あり(自由質問)		処置群	41	27
(e) 再話あり(質問なし)	処置群	対照群	36	75

注. 条件(a)は, 文再認課題の結果に再話が影響しないという意味で「再話なし条件」と名付けた。条件(b)～(d)は, 第2～4週の3回の授業において, 一回でも, 新たなIUの再生につながる足場がけの質問をした協力者数を示す。一方で, もともと条件(b)～(d)に割り当てられていた協力者の中で, 足場がけの質問をしなかった協力者は, 条件(e)に再編されている。

再話あり(質問なし)の条件の協力者数が多くなっているのは, 前述の通り, 足場がけの質問を行う条件で, 協力者全員が, 質問の支援を行ったわけではないことに起因する。実際, 授業者が用意した事実質問と推論質問はほとんど使用されなかった。具体的に, 条件(b)では, IUの再生に寄与する事実質問を行った協力者は37名中7名であった。条件(c)では, 推論質問を行い, IU再生に寄与した協力者は41名中わずか2名であった。条件(d)でIUの再生に寄与する自由質問を行った協力者は41名中27名であった。

以上の結果から, 単純にクラス別の5条件を比較することは, 分析手順として適切でないと判断し, 今後の分析は, 実際に各条件で取り組んだ人数に基づいて条件間比較を行うこととした。また, 条件(b)と条件(c)はデータ数が少ないため, 条件(d)とまとめて, 再話あり(質問あり)条件とした。結果として, 再話あり(質問なし)条件(n = 75), 再話あり(質問あり)条件(n =7+2+27 = 36), 再話なし条件(n = 29)の3条件に再編した。

3.4.1 IU再生率

収集された全ての再話プロトコル(再話者が再話で発話した内容)について, 著者が文字起こしを行った。そして, 著者を含めた2名の評価者で, IU再生率を算出した。まず全体の15%について, 2名の評価者が独立して IUの再生率を求めた。その後, 評価者間一致率を測定した (FASHION; 96.59%, ROCKET; 95.67%, HATS; 92.06%, AUSTRALIA; 92.67%, PIGS; 94.12%)。十分に高い一致率が得られたため, 疑義が生じた箇所について議論を経て解決し, 残りの再話プロトコルについては, 著者が一人で IU再生率を算出した。

3.4.2 発話流暢性

第1週と第5週の再話プロトコルを分析して, 発話流暢性の指標を算出した。流暢性の指標として8つの指標を用意した(表5)。無声ポーズと判断する基準は先行研究によって異なるが, 本研究では500ミリ秒よりも長い沈黙を無声ポーズとした(Maeda, 2021)。有声フィラーと無声ポーズはまとめてポーズとして扱った(Ogawa, 2022)。有声フィラーは, 日本語と英語の区別をしなかった。言い間違え(false start), 言い換え(reformulation), 繰り返し(repetition)の3つを言いよどみ(disfluency markers)に分類した(Ogawa, 2022)。分析の際は, これら3つの合計数を言いよどみ数とした。無声ポーズ, 有声フィラーは著者1名の判断で分類したが, 言いよどみについては, 著者を含めた2名の評価者で, 再話プロトコルの15% を別々に評価して, 評価者間一致率を求めた(FASHION; 98.16%, PIGS; 98.86%)。言いよどみの識別について十分に高い一致率が得られたため, 疑義が生じた箇所については議論を経て解決し, 残りの再話プロトコルについては, 著者一人で言いよどみの時間を測定した。無声ポーズ, 有声フィラー, 言いよどみに要した時間は音声分析ソフトウェアPraat(Boersma & Weenink, 2009)を用いて測定した。総シラブル数と総単語数は Syllable Counter(Arczis Web Technologies, Inc., 2023)を用いて計算した。

その他の指標については，表5に示した方法で算出した。

■表5: 英語による発話流暢性の指標

	流暢性の指標	流暢性の分類	計算方法
1	総シラブル数 (number of syllables)	speed fluency	有声フィラーと言いよどみを除いた発話データをもとに，Syllable Counter を用いて計算
2	1秒間の平均シラブル数 (syllables per second)	speed fluency	総シラブル数 / [総発話時間 − (無声ポーズ時間 + 有声フィラー時間 + 言いよどみ時間)]
3	総単語数 (number of words)	speed fluency	有声フィラーと言いよどみを除いた発話データをもとに，Syllable Counterを用いて計算
4	1分間の平均単語数 (words per minute)	speed fluency	{総単語数 / [総発話時間 − (無声ポーズ時間 + 有声フィラー時間 + 言いよどみ時間)]} * 60
5	ポーズ間の平均シラブル数 (mean length of run)	speed fluency	総シラブル数 / [(無声ポーズ数 + 有声フィラー数) + 1]
6	平均ポーズ時間 (秒) (mean pause length)	breakdown fluency	(無声ポーズ時間 + 有声フィラー時間) / (無声ポーズ数 + 有声フィラー数)
7	発話時間の割合 (%) (phonation time ratio)	breakdown fluency	{[総発話時間 − (無声ポーズ時間 + 有声フィラー時間 + 言いよどみ時間)] / 総発話時間} * 100
8	1分間の平均言いよどみ数 (repairs per minute)	repair fluency	{言いよどみ数 / [総発話時間 − (無声ポーズ時間 + 有声フィラー時間)]} * 60

3.4.3 文再認課題

邑本 (2000) に従い，7段階 (0点〜6点) で協力者の再認確信度を評価した。最初に，テスト項目の文が登場したか，登場していないかを回答させ (二択)，次に，その確信度を4段階で回答させた (全く確信がない・あまり確信がない・まあまあ確信がある・非常に確信がある)。「全く確信がない」という回答は「登場した」，「登場していない」の回答に関わらず3点とした。したがって再認確信度評定値は，テスト項目が登場したと判断される場合は，4〜6点のいずれかの値となり，登場しなかったと判断される場合は，0〜2点のいずれかの値となる。いずれとも判断がつかない場合には3点となる。この3点を基準にして，0点は，確信を持ってテスト項目がテキストに登場しなかったと判断したことを示し，6点は，確信を持ってテスト項目がテキストに登場したと判断したことを示す。

第2週〜第4週の文再認課題のデータを使用し，協力者ごとに，言い換え文，推論文，錯乱文の再認確信度評定値を収集した。第1週目と第5週目のデータは，すべての協力者で再話条件，質問条件いずれも同条件で収集されており研究課題の検証には直接関係しないため，分析の対象としなかった。

4 結果

4.1 読解熟達度テスト

信頼性係数 (クロンバックα) は $\alpha = .77$ だった。新たに再編成した3条件ごとの読解熟達度テストの結果は，表6の通りである。

■表6: 3条件ごとの読解熟達度テストの記述統計量

条件	対照群・処置群		人数	平均	標準偏差	最小値	最大値
	研究課題2	研究課題3					
(A) 再話なし	対照群		29	10.86	3.16	2	16
(B) 再話あり(質問あり)		処置群	36	9.47	3.49	5	19
(C) 再話あり(質問なし)	処置群	対照群	75	11.96	4.39	4	26

上記の3条件間で, 読解熟達度テストの得点に有意な差がないかどうかを検証した。クラスカル・ウォリス検定の結果, 3条件の間に有意な差があることが示された($H(2) = 10.85$, $p = .004$)。多重比較の結果, 条件(B)が条件(C)よりも有意に熟達度が低いことが示された($p = .003$)。研究課題3の検証のためには, 条件(B)と条件(C)を比較する必要があるが, このままでは, 協力者の熟達度が統制できないことになる。前述の通り, 条件(B)の協力者は, 第2〜4回の複数の授業において, 一度でも足場がけの支援を受けた協力者であり, つまり, 足場がけの支援を受けた回と支援を受けていない回が混在している。そこで研究課題3の検証では, 条件(B)の協力者内で, 足場がけの質問があったデータとなかったデータとを比較, 分析することにした。今後の分析において, 協力者の英語熟達度の要因を考慮する場合, 読解熟達度テストの得点は z得点に換算して使用した。

4.2 IU再生率

最初に, 条件(C)の IU再生率を算出した(表7)。次に, 条件(B)の中で, 足場がけの質問の支援があった条件(B-1)となかった条件(B-2)とでそれぞれの IU再生率を算出した(表8, 9)。これらの結果は, 質問条件に関わらず, 読解した内容の半分程度は再話をすることができたことを示しており, テキストの難易度は適切であったと考えられる。

■表7: 条件(C)(再話あり (質問なし) 条件)におけるIU再生率の記述統計量

週	テキスト	人数	平均	標準偏差	95%信頼区間	最小値	最大値
第1週	FASHION	75	52.84	19.66	[48.32, 57.37]	15.79	100.00
第2週	ROCKET	64	50.00	21.04	[44.74, 56.26]	10.00	100.00
第3週	HATS	71	54.19	19.71	[49.53, 58.86]	14.29	95.24
第4週	AUSTRALIA	63	50.95	17.39	[46.57, 55.33]	25.00	100.00
第5週	PIGS	75	58.72	18.45	[54.48, 62.97]	21.74	100.00

■表8: 条件(B-1)(再話あり (質問あり) 条件)におけるIU再生率の記述統計量

週	テキスト	人数	平均	標準偏差	95%信頼区間	最小値	最大値
第2週	ROCKET	13	48.08	19.32	[36.40, 59.75]	25.00	95.00
第3週	HATS	15	61.31	16.02	[52.77, 69.85]	28.57	85.71
第4週	AUSTRALIA	19	47.11	16.69	[39.06, 55.15]	25.00	95.00

■表9: 条件(B-2)(再話あり (質問なし) 条件)におけるIU再生率の記述統計量

週	テキスト	人数	平均	標準偏差	95% 信頼区間	最小値	最大値
第1週	FASHION	36	54.09	15.18	[48.96, 59.23]	21.05	84.21
第2週	ROCKET	21	46.43	20.32	[37.18, 55.68]	20.00	90.00
第3週	HATS	18	46.56	12.17	[40.51, 52.61]	33.33	71.43
第4週	AUSTRALIA	16	47.81	13.16	[40.80, 54.83]	20.00	70.00
第5週	PIGS	36	56.04	16.39	[50.49, 61.59]	30.43	100.00

足場がけの質問の効果を検証するために, 条件(B-1)と条件(B-2)を比較し, IU再生率に違いがあるのか検証した(資料2)。線形混合モデルを使用し, 目的変数を IU再生率, 説明変数のうち, 固定効果は, 質問条件(あり・なし), 協力者の英語熟達度(読解熟達度テストのz得点), 使用テキスト(ROCKET・HATS・AUSTRALIA)とし, 質問条件と使用テキストの交互作用項も設定した。また変量効果として, 協力者をランダム切片に, 協力者の英語熟達度をランダム傾きに設定した。最適モデルを推定した結果, 質問の有無, 熟達度, 使用テキストの主効果が有意であった。これは, 足場がけの質問がある方が IU再生率が高くなること, 熟達度の高い協力者ほど IU再生率が高くなること, IU再生率はテキストによって異なることを意味する。交互作用は確認できなかった。使用テキストについては, 多重比較の結果, HATS の IU再生率が AUSTRALIA の IU再生率よりも有意に高かった(β = 6.29, SE = 2.57, t = 2.45, p = .044)。分析の結果から, 足場がけの質問を行うことが, 滞った再話の再開を促し, 結果としてより多くのテキスト情報を再生する支援になることが示された。また, 足場がけの質問の効果にはテキスト要因も関係することが示唆された。

4.3 発話流暢性

発話流暢性については, 質問なし条件で再話に取り組んだ協力者(条件(C):n = 75)を分析した。これは, 再話における足場がけの質問の効果を取り除いたうえで, 流暢性の指標の変化を分析するためである。流暢性の指標についての記述統計量は表10の通りである。

■表10: 流暢性の指標についての記述統計量 (第1週及び第5週)

流暢性の指標		第1週				第5週			
		データ数	平均	標準偏差	95% 信頼区間	データ数	平均	標準偏差	95% 信頼区間
1	総シラブル数(number of syllables)	75	93.48	29.76	[86.63, 100.33]	75	99.60	30.45	[92.60, 106.60]
2	1秒間の平均シラブル数(syllables per second)	75	2.51	0.36	[2.43, 2.59]	75	2.57	0.35	[2.49, 2.65]
3	総単語数(number of words)	75	61.19	21.41	[56.26, 66.11]	75	70.25	22.14	[65.16, 75.35]
4	1分間の平均単語数(words per minute)	75	97.82	16.62	[93.99, 101.64]	75	108.92	16.45	[105.14, 112.71]
5	ポーズ間の平均シラブル数(mean length of run)	75	4.04	1.42	[3.72, 4.37]	75	4.00	1.45	[3.66, 4.33]
6	平均ポーズ時間(秒)(mean pause length)	75	1.83	0.86	[1.63, 2.03]	75	1.54	0.44	[1.44, 1.64]
7	発話時間の割合(%)(phonation time ratio)	75	46.45	12.62	[43.55, 49.36]	75	48.62	11.40	[45.99, 51.24]
8	1分間の平均言いよどみ数(repairs per minute)	75	8.03	9.44	[6.54, 9.51]	75	7.66	6.53	[6.15, 9.16]

研究課題1を検証するために，線形混合モデルを用いて分析した。固定効果は，週の推移（第1週・第5週），協力者の英語熟達度（読解熟達度テストのz得点）とし，変量効果は，協力者をランダム切片とした。ランダム傾きはモデルが収束しないため，変量効果に含めなかった。各指標の最適モデルと推定値は資料3の通りである。各流暢性の指標について，有意差の有無を表11にまとめる。

■表11: 発話流暢性について有意差が認められた指標

	流暢性の指標	流暢性の分類	週の推移	熟達度
1	総シラブル数（number of syllables）	speed fluency	有意差あり	n.s.
2	1秒間の平均シラブル数（syllables per second）	speed fluency	n.s.	有意差あり
3	総単語数（number of words）	speed fluency	有意差あり	有意差あり
4	1分間の平均単語数（words per minute）	speed fluency	有意差あり	有意差あり
5	ポーズ間の平均シラブル数（mean length of run）	speed fluency	n.s.	n.s.
6	平均ポーズ時間（秒）（mean pause length）	breakdown fluency	有意差あり	有意差あり
7	発話時間の割合（%）（phonation time ratio）	breakdown fluency	n.s.	有意差あり
8	1分間の平均言いよどみ数（repairs per minute）	repair fluency	n.s.	n.s.

注. n.s. = 有意差なし。

5週にわたる授業に参加することで，有意に向上した指標は，総シラブル数，総単語数，1分間の平均単語数，平均ポーズ時間の4つであり，総シラブル数を除いて，熟達度の高い学習者の方が有意に高い指標を示した。repair fluency に関する指標は，本研究の前後で有意な変化は認められなかった。

4.4 文再認課題

第2週～第4週の文再認課題の信頼性係数を算出した。一つのテキストにつき，3種類の文再認課題を作成したため，合計9つの課題ごとに信頼性係数を算出した。文再認課題の信頼性係数の平均値は α = .60だった。

研究課題2を検証するために，4.1, 4.2で設定した条件ごとに（条件（A），条件（B-1），条件（B-2），条件（C）），再認確信度評定値を算出した。各条件の再認確信度評定値の記述統計量は，表12, 13が示す通りである。

■表12: 再話条件 (あり・なし) ごとの再認確信度評定値の記述統計量

条件	テスト項目	項目数	平均	標準偏差	95% 信頼区間	最小値	最大値
(A) 再話あり	言い換え文	659	4.95	1.54	[4.83, 5.07]	0	6
	推論文	577	3.71	2.14	[3.54, 3.89]	0	6
	錯乱文	490	0.43	0.85	[0.35, 0.50]	0	6
(C) 再話なし	言い換え文	266	4.64	1.61	[4.45, 4.84]	0	6
	推論文	246	3.89	1.93	[3.65, 4.13]	0	6
	錯乱文	194	0.87	1.13	[0.71, 1.03]	0	6

注. 条件(A) は, 足場がけの質問を行わない再話条件である。

■表13: 質問条件 (あり・なし) ごとの再認確信度評定値の記述統計量

条件	テスト項目	項目数	平均	標準偏差	95% 信頼区間	最小値	最大値
(B-1) 質問あり	言い換え文	170	4.93	1.58	[4.68, 5.18]	0	6
	推論文	157	3.87	2.14	[3.52, 4.22]	0	6
	錯乱文	123	0.41	0.88	[0.24, 0.58]	0	6
(B-2) 質問なし	言い換え文	160	4.85	1.51	[4.62, 5.08]	0	6
	推論文	144	3.52	2.01	[3.19, 3.84]	0	6
	錯乱文	110	0.56	0.83	[0.41, 0.71]	0	4

4.4.1 英語による再話が内容理解に及ぼす影響

研究課題2を検証するために, 条件(A)と条件(C)を比較して, 再話条件が3つのテスト項目の再認確信度に影響を与えるのかについて分析した。線形混合モデルを使用し, 目的変数は再認確信度評定値とし, 説明変数は, テスト項目(言い換え文・推論文・錯乱文), 再話条件(あり・なし), 協力者の英語熟達度(読解熟達度テストのz得点)として, これらの説明変数を固定効果とした。さらにテスト項目と再話条件の交互作用項も固定効果に含めた。また変量効果として, 協力者をランダム切片に, 協力者の英語熟達度をランダム傾きに設定した。モデルを検討した結果, 最適なモデルの推定値は, 資料4の通りである。

分析の結果, テスト項目と再話条件の交互作用が有意であったため, 単純主効果の検定を行った。テスト項目については多重比較の結果, 再話条件を問わず, 再認確信度評定値は, 錯乱文 < 推論文 < 言い換え文の順で有意に高かった(すべての多重比較において $p < .001$)。次にテスト項目ごとに, 再話の効果を検証した結果, 再話を行うことで, 錯乱文の再認確信度評定値が有意に減少し($\beta = 0.41$, $SE = 0.15$, $t = 2.71$, $p = .007$), また言い換え文の再認確信度評定値が有意に増加することが示された($\beta = -0.33$, $SE = 0.13$, $t = -2.43$, $p = .016$)。一方で, 推論文については英語による再話の効果が認められなかった($\beta = 0.15$, $SE = 0.14$, $t = 1.10$, $p = .274$)。

4.4.2 足場がけの質問が内容理解に及ぼす影響

次に, 研究課題3を検証するために, 条件(B-1)と条件(B-2)を比較して, 足場がけの質問の有無によって, 各テスト項目の再認確信度評定値が異なるのかについて検証した。線形混合モデルを使用した結果, いず

れのモデルもモデルの前提が満たされなかったため，線形混合モデルによる分析を断念せざるを得なかった。そこで，再認確信度評定値を7値（0点〜6点）から2値（0点，1点）に変換して，一般化線形混合モデルを使用して再度分析を試みた。テスト項目が「登場した」と判断した場合の再認確信度評定値（4〜6点）を1点，「登場しない」と判断した場合の再認確信度評定値（0〜2点）を0点に変換した。判断がつかなかった回答（3点）は削除した（削除した項目数：n = 16）。目的変数に再認確信度評定値を設定し，説明変数に，テスト項目（言い換え文・推論文・錯乱文），質問条件（あり・なし），協力者の英語熟達度（読解熟達度テストのz得点）を設定し，これらを固定効果とした。テスト項目と質問条件の交互作用項も固定効果に設定した。glmer関数内のfamily指定はbinomialとし，リンク関数はlogitとした。また変量効果として，協力者をランダム切片に，協力者の英語熟達度をランダム傾きに設定した。モデルを検討した結果，最適なモデルの推定値は，資料5の通りである。質問条件を含まないモデルが最適モデルと判断されたため，再話が滞った際の足場がけの質問が再認確信度に影響を及ぼすわけではないことが示された。よって，足場がけの質問が内容理解に及ぼす影響は，テキストベース，状況モデルの両レベルにおいて確認できなかった。

5 考察

5.1 英語による再話が発話流暢性に及ぼす影響

分析の結果，speed fluencyについては，総シラブル数と総単語数が有意に増加したことから，英語による再話を授業に取り入れることで，読解した内容について英語で話す量が増加する可能性が示唆された。これは佐々木（2020）と一致する結果となった。スピーキング能力の高い学習者の発話量は，そうでない学習者と比較して多くなるので（小泉・栗嵜，2002），英語による再話は，speed fluency向上に一定の効果が認められると言える。しかし表10が示すように，総単語数の増加数（約9語の増加）と比べて総シラブル数の増加数（約6シラブルの増加）が明らかに少ないため，speed fluencyが向上するかどうかは，テキスト要因も関係する可能性がある。可能性の一つとして，第5週のテキスト（PIGS）の再話では，シラブル数が少なく短い英単語が多く使用されたのかもしれない。1分間の平均単語数が増加した一方で，1秒間の平均シラブル数が変化しなかったことも，この可能性を示唆していると思われる。また総シラブル数が有意に増加したとはいえ，資料3が示すようにMarginal R^2の値が極めて小さく，個人差が大きい点にも注意が必要である。少なくとも本研究のような5週にわたる期間では，英語による再話がspeed fluencyに及ぼす効果は，発話スピードよりも発話量に対して顕著であり，学習者の個人差やテキストで使用される語彙の種類といったテキスト要因により変化する可能性がある。また，ポーズ間の平均シラブル数が増加しなかったことから，英語の発話量が増えても，それに応じてポーズ数も増加したと考えられる。これは定型表現（formulaic expressions）の使用状況と関係した可能性がある（e.g., Kormos, 2006）。定型表現の習得が進めば，特に初級学習者に多く見られる単語間の不必要なポーズは減少し，ポーズ間の平均シラブル数の点でも speed fluency は向上する可能性がある。

次に breakdown fluency については，平均ポーズ時間は有意に減少し，また熟達度が高い学習者の方が，有意に平均ポーズ時間が短いことが明らかになった。よって，英語による再話に継続して取り組むことで breakdown fluency が向上する可能性が示された。平均ポーズ時間の減少は，継続して再話に取り組むことによって，英語で発話する際に必要となる知識が，宣言的知識から手続き的知識へと変容したことに起因すると考えられる（Skehan, 1998）。

最後にrepair fluencyについては，1分間の平均言いよどみ数は，5週に及ぶ再話を取り入れた授業を行うことでは減少しなかった。再話がspeed fluencyとbreakdown fluencyの向上に一定の効果を示しながら，repair fluencyは変化しなかったということは，単語を固まりでとらえてフレーズで使用する能力の向上には，speed fluencyや breakdown fluency の向上よりも時間を要するのかもしれない。

5.2 英語による再話が内容理解に及ぼす影響

英語による再話を行うことで，言い換え文の再認確信度評定値が有意に高まり，錯乱文の再認確信度評定値が有意に減少した。一方で，推論文については，英語による再話の効果が認められなかった。つまり，英語による再話は，テキストベースレベルの表象を強化し，その結果，テキストの内容にそぐわない情報に気づく可能性を高めることが示唆された。Ito（in press）は，日本語による再話を行うと，言い換え文と推論文の両方の再認確信度評定値が高まる可能性を示しているが，英語による再話と日本語による再話では，各理解レベルに及ぼす影響も異なる可能性がある。本研究で，言い換え文の再認確信度評定値が高まった理由として，テスト効果が現れた可能性がある（Roediger et al., 2011）。英語による再話は，多くの場合，テキストに登場した表現を用いて行う。学習者はテキストで使用された表現を想起し，それらを使用することで，使用した表現に対する記憶が強化され，結果としてテキストベースの強化につながったと考えられる。錯乱文に対する再認確信度評定値が減少した理由も同様で，再話によりテキストベースが強化された結果，テキストの内容にそぐわない情報により気づくことができるようになったと考えられる。

一方で，推論文については英語による再話の効果が認められなかったため，英語による再話がより豊かな状況モデルの構築には必ずしもつながらないことが示された。これは，初級英語学習者ほど，テキストの表現に依存して再話を行う可能性が高いため，表現形式により注意を向けた結果，推論生成に認知資源を割く余裕がなかった可能性がある。伊藤・垣花（2009）は，日本語による内容説明の中で推論付加が認められたと報告しているが，本研究のような英語による再話では，協力者は表現形式により注意を向けて，結果として情報間のつながりを推論で補ったり，背景知識とテキスト情報を統合したりといった処理に至らなかった可能性がある。一方で，再話によりテキストベースが強化された結果，明示的に示されていないが妥当な推論から導かれる情報を含んだ推論文を，テキストの内容と一致しないと判断した可能性もある（邑本, 2000）。いずれにせよ，英語による再話は，学習者が読解時点で構築した状況モデルをより精緻化するというよりも，テキストベース強化につながる可能性が高いことが示唆された。

5.3 足場がけの質問が内容理解に及ぼす影響

本研究では，足場がけの質問による支援が必ずしも内容理解に影響を及ぼすわけではないことが示唆された。具体的に，テキストベースレベル，状況モデルレベルのいずれにおいても，足場がけの質問の効果は確認できなかった。これは，本研究で使用された質問の多くが「自由質問」で，再話が滞った際の支援が単語レベルの再生にとどまったことが原因だと考えられる。また，多くの場合，そもそも質問の形をとらず，再生できない単語に気づかせようとして，まずは日本語でその単語の意味を伝えて，それでうまくいかなければ，英語で該当する単語そのものを伝えて支援していた。（例1）では，再話者が，宇宙ゴミ（space junk）がバラバラになって地球に落下するということを伝えようとして，再話が滞っている。この時に聞き手は，落下する場所そのもの（Global South）を示している。また，more likely to hitという表現の hit が出てこない場面で，そのまま hit を示している。（例2）も同様に，想起できない単語（fall down）を示すことで再話を支援している。よって本研究では，足場がけの質問は，テキスト情報の再生率を高めることにつながったとしても単語レベルの情報の想起にとどまり，再話者に読解した内容について深い処理を促すことにはつながらなかったと考えられる。また全体として，質問による支援の回数が少なく，1回の再話において滞った箇所一つに対して足場がけの質問を行って，そこで支援が終わることが多かった。支援の回数が少なかったことも，足場がけの質問が内容理解に影響を与えるまでに至らなかった要因かもしれない。

（例 1）

再話者：Have you ever worried about hit, a rocket hitting you? ... They then break up and fall down, fall, fall back to earth.

聞き手：南半球の発展途上国が出てくる．Global South.

再話者：Global South. Likely, likely, more like, more likely to, to.

聞き手：*Hit.*

再話者：To hit, to hit Global South.

（例 2）

再話者：Have you ever worried, worries hitting a rocket? ... But many break, sixteen percent, sixty percent rocket is break. So, so.

聞き手：*And* 落ちる．

再話者：And falling, fall, fall, fa, falling.

聞き手：*Fall down.*

再話者：Fall down of earth, of earth. So, so rocket company risk.

6 結論・教育的示唆・今後の課題

本研究は，英語による再話を授業に取り入れた場合に，(a)発話流暢性が向上するのか，(b)英語による再話が内容理解に影響を及ぼすのか，(c)どのような足場がけの質問が再話者の状況モデル構築につながるのかを検証した。5週にわたる授業に取り組んだ結果，speed fluency と breakdown fluency については，一定程度の向上につながることが示された。また英語による再話により，学習者はテキストの表現形式そのものに注意を向けるようになり，テキスト表現に忠実な情報の記憶が強化されることで読解時に構築したテキストベースが強化される可能性が示唆された。結果として，テキストの内容にそぐわない情報に気づく可能性が高まるが，一方で，推論情報や背景知識とテキスト情報が結びついたより精緻な状況モデルの構築には必ずしもつながらない可能性が示唆された。最後に，再話が滞った際の足場がけの質問について，聞き手は，授業者が用意した質問よりも，再話者に語彙レベルの知識の想起を促す質問を多く行った。この支援により，再話者は滞りの原因となる箇所に語彙レベルで気づくことで，さらに再話を続けることが可能になり，結果としてテキスト情報をより多く再生することができた。一方で，本研究で学習者が使用した足場がけの質問は，支援の結果として発話量が増えたとしても，それがテキストベースレベル，状況モデルレベルの理解を深めるわけではないことが示唆された。

本研究の結果から，英語による再話を行うことで，流暢性の向上と内容理解を深めるという2つの目的を同時に達成し得る可能性が示された。しかし，内容理解をテキストベースレベルと状況モデルレベルそれぞれで考えた場合，英語による再話に取り組む際に，よりテキスト表現に依存する可能性が高い英語初級学習者に対しては，再話が内容理解に及ぼす効果は，テキストベースレベルの表象の強化にとどまる可能性がある。よって，英語による再話の際には，テキストの難易度を考慮したり，十分な準備段階を設けたりといった配慮がより求められるであろう。より伝えたい内容に意識を向けた再話ができれば，より豊

かな状況モデル構築につながると思われる。足場がけの質問については，授業者が用意した質問は，学習者にとって利用することが難しく，特に推論質問についてはこの傾向が顕著であった。これは，本研究の協力者が英語初級学習者（CEFRレベルでA1～A2レベルに相当）であったため，再話が滞った際に適切な足場がけの質問を行うことが，そもそも難しかったことも一因であろう。学習者が自由に質問をして再話の支援を行う場合は，足場がけの質問が局所的な処理しか引き起こさず，より精緻な状況モデル構築につながらない可能性があり，これは特に英語初級学習者に当てはまることだと思われる。よって，どのような質問をすれば，再話者がより深く自分の理解を振り返ることにつながる支援となるのか，また滞った再話を進めることにつながるのかについて，学習者自身が考える必要性も指摘しておきたい。

本研究の限界点として，再話条件，質問条件を厳密に統制することができなかったことが挙げられる。特に，再話条件については，再話を行った群では，再話を行わなかった群よりも，そもそもテキストに触れる時間が多く設けられていたため，この点が内容理解（再認確信度）に与えた影響は否定できない。今後は，再話を行わない代わりに音読を行う条件を設定する等，テキストに触れる時間を統制したうえで，再話の効果を検証する必要がある。また，本研究で使用した文再認課題の信頼性係数が高くなかったため，より信頼性の高いテストを作成したうえで，本研究と同様の結果が得られるのかどうか，今後検証が必要であると思われる。

謝辞

本研究を実施する貴重な機会を与えてくださいました公益財団法人 日本英語検定協会の皆様，ならびに選考委員の先生方にこの場をお借りして厚く御礼申し上げます。中でも，研究助言者の西垣知佳子先生には貴重なご助言と的確なご指導をいただき，深く感謝申し上げます。また，筑波大学大学院の卯城祐司先生には，研究実施から執筆にいたるまで，多くのご指導やお言葉をいただき，深く感謝申し上げます。さらに，同期の小室竜也さんをはじめ，後輩の皆様にも多くのアドバイスやサポートをしていただき，深く感謝申し上げます。最後になりますが，実験協力者の学生をはじめとして本研究の実施にご協力いただいたすべての皆様に心より御礼申し上げます。

注

1. 伊藤・垣花（2009）は「再話」という言葉は使用していないが，他者に向けた説明が話者自身の理解に効果を及ぼすのかという研究課題を検証したものであり，再話にも適用できると考える。
2. 伊藤・垣花（2009）は「意味付与的説明」に加えて，この意味付与的説明の「繰り返し」も内容理解に及ぼす効果の要因であるとしている。
3. 清水（2005）は推論質問をさらに，（a）因果的先行詞，（b）上位ゴール，（c）登場人物の感情的反応，（d）因果的結末，（e）道具，（f）下位ゴール／行為，（g）状態，（h）読み手の感情の8種類に下位分類している。
4. 実際には，第1週目の授業の前の週に，読解熟達度テストと，英語による再話を取り入れた授業により学習者の読解に対する態度や方略がどのように変化するのかについて調査するアンケートを実施している。また，同様のアンケートを事後アンケートとして第5週目の授業の次の週に実施している。よって，本来は7週にわたる調査になるのだが，このアンケートに関するデータは，本研究の研究課題と直接の関係がないため，ここでは取り扱わない。
5. T/Fクエスチョンは，足場がけの質問とは異なる質問とした。
6. 第1週目と第5週目の比較は，再話プロトコルの分析に基づく発話流暢性の比較のため，文再認課題のデータは分析していない。
7. 本研究では，再話条件（再話あり・再話なし）の比較をする際に，学生が取り組む活動にできるだけ差をつけないという教育的配慮から，文再認課題に取り組むタイミングを操作することで，再話条件を設定した。具体的には，一つのクラスで再話と文再認課題の順序を入れ替えるという対応を行った。

引用文献

Arczis Web Technologies, Inc. (2023). Syllable count. https://www.syllablecount.com/

Banville, S. (2021, March 21). *Pigs are intelligent enough to play video games.* Breaking News English. https://breakingnewsenglish.com/2103/210321-pigs-0.html

Banville, S. (2022, February 10). *Fast fashion is harming the environment.* Breaking News English. https://breakingnewsenglish.com/2202/220210-fast-fashion-0.html

Banville, S. (2022, July 14). *Bigger risk of being killed by falling space rocket.* Breaking News English. https://breakingnewsenglish.com/2207/220714-falling-rockets-0.html

Banville, S. (2022, July 21). *Australia's 'shocking' decline in environment.* Breaking News English. https://breakingnewsenglish.com/2207/220721-biodiversity-loss-0.html

Banville, S. (2022, October 27). *Anti-cheating exam hats used in Philippines college.* Breaking News English. https://breakingnewsenglish.com/2210/221027-anti-cheating-hats-0.html

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer. https://www.fon.hum.uva.nl/praat/

Foster, P. (2020). Oral fluency in a second language: A research agenda for the next ten years. *Language Teaching, 53,* 446–461.

Gambrell, L. B., Koskinen, P. S., & Kapinus, B. A. (1991). Retelling and the reading comprehension of proficient and less-proficient readers. *The Journal of Educational Research, 84*(6), 356–362.

語学教育研究所 (2014).『英語指導技術ガイドQ&A－授業の悩みにこたえる26のレシピー』東京: 開拓社.

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I., Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1–20). Amsterdam: John Benjamins Publishing Company.

Ikeno, O. (1996). The effects of text-structure-guiding questions on comprehension of texts with varying linguistic difficulties. *JACET Bulletin, 27,* 51–68.

Ito, S. (in press). A two-step approach to measuring the effects of retelling on mental representations in Japanese EFL learners: An analysis using sentence recognition tasks. *JLTA Journal*.

伊藤貴昭・垣花真一郎 (2009).「説明はなぜ話者自身の理解を促すか－聞き手の有無が与える影響－」『教育心理学研究』第57号, 86-98.

Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics, 19*(3). 357-377.

Kai, A. (2008). The effects of retelling on narrative comprehension: Focusing on learners' L2 proficiency and the importance of text information. *ARELE, 19,* 21–30.

Kai, A. (2009). Achieving global coherence through retelling. *ARELE, 20,* 41–50.

Karimi, M. N., & Richter, T. (2023). Belief-biased representations of textual information in bilinguals: Language as a source characteristic. *Current Psychology, 42,* 9852-9866.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294–303.

Kintsch, W. (2013). Revisiting the construction-integration model of text comprehension and its implications for instruction. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 807–839). Newark: International Reading Association.

Kintsch, W. & Kintsch, E. (2005), Comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 71–92). Mahwah: Center for Improvement of Early Reading Achievement.

小泉利恵・栗嵜逸美 (2002).「日本人中学生のモノローグにおけるスピーキングの特徴」『関東甲信越英語教育学会研究紀要』第16号, 17-28.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah: Lawrence Erlbaum Associates Publishers.

Koskinen, P. S., Gambrell, L. B., Kapinus, B. A., & Heathington, B. S. (1988). Retelling: A strategy for enhancing students' reading comprehension. *The Reading Teacher, 41*(9), 892–896.

Maeda, H. (2021). Speaking performance among high school students at different CEFR levels: A comparative study on the TEAP speaking section. *ARELE, 32,* 33-48.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1–43.

Morrow, L. M. (1985). Retelling stories: A strategy for improving young children's comprehension, concept of story structure, and oral language complexity. *The Elementary School Journal, 85*(5), 647–661.

邑本俊亮 (2000).「第二言語の文章理解過程に及ぼす習熟度の影響 －文再認課題による実験－」『読書科学』第42(2)号, 43-50.

Ogawa, C. (2022). CAF indices and human ratings of oral performances in an opinion-based monologue task. *Language Testing in Asia, 12*(4).

引用文献

旺文社 (2010a).『英検準2級過去6回全問題集』東京：旺文社.

旺文社 (2010b).『英検2級過去6回全問題集』東京：旺文社.

小河原義朗・木谷直之 (2020).『「再話」を取り入れた日本語授業 初中級からの読解ー読んで理解したことが伝えられるようになるためにー』東京：凡人社.

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Roediger, H. L. III., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Cognition in education* (pp. 1-36). Cambridge: Elsevier Academic Press.

佐々木啓成 (2020).『リテリングを活用した英語指導ー理解した内容を自分の言葉で発信する』東京：大修館書店.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory and Language, 25*(3), 279-294.

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching, 54*(2), 79-95.

清水真紀 (2005).「リーディングテストにおける質問タイプーパラフレーズ・推論・テーマ質問と処理レベルの観点からー」『STEP BULLETIN』第17号, 48-62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

鷲見俊幸 (2015).「英語の授業における再話の効果」『中部地区英語教育学会紀要』第44号, 169-174.

Tachibana, K. (2016). Cooperative retelling with Japanese high school EFL learners: Facilitating improvement in English language proficiency. *Journal of the Chubu English Language Education Society, 45,* 103-110.

田中武夫・辻智生 (2015).「推論発問および評価発問を活用した英語リーディング指導の実践ー高等学校における1年間の実践事例を通してー」『教育実践学研究』第20号, 159-171.

卯城祐司・甲斐あかり・小泉利恵・清水遥・清水真紀・中川知佳子・星野由子 (2009).『英語リーディングの科学ー「読めたつもり」の謎を解く』東京：研究社.

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes, 54*(5-6), 360-372.

Wilson, R. M., Gambrell, L. B., & Pfeiffer, W. R. (1985). The effects of retelling upon reading comprehension and recall of text information. *The Journal of Educational Research, 78*(4), 216-220.

Wolfe, M. B. W., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology, 80*(3), 341-362.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162-185.

資料1：用意した事実質問と推論質問

（第 2 週）Bigger risk of being killed by falling space rocket

事実質問

Q1. What do we have to worry about?
Q2. What is space crowded with?
Q3. What do the scientists want rocket companies to do?
Q4. What risk did scientists look at?
Q5. How many rockets went into space in 2021?
Q6. Where is space junk likely to hit?

推論質問

Q1. Why do we have to worry about a rocket hitting us?
Q2. Why could a piece of space junk kill someone?
Q3. Why is space crowded with space junk?
Q4. Why do rocket companies have to control how space junk comes back to Earth?
Q5. Why do scientists think about the risk of a rocket falling someone?
Q6. Why are rocket companies putting poorer nations at risk?

（第 3 週）Anti-cheating exam hats used in Philippines college

事実質問

Q1. What do schools in the Philippines try hard to do in exams?
Q2. What must students in the Philippines wear in exams?
Q3. How do students feel when they make their own hats?
Q4. What have gone viral on social media?
Q5. What kind of things do students use to make anti-cheating hats?
Q6. Who was a professor proud of?

推論質問

Q1. Why must students in the Philippines wear hats in exams?
Q2. Why are the students happy when they make their hats?
Q3. Why do the students prepare cardboard boxes, egg cartons and coat hangers?
Q4. Why have photos of the hats gone viral on social media?
Q5. Why did many students finish the tests early?

（第 4 週）Australia’s ‘shocking’ decline in environment

事実質問

Q1. What have scientists reported?

Q2. What has been lost in Australia?

Q3. What is damaging Australia’s nature?

Q4. What does Australia make every five years?

Q5. What did a nature group say?

Q6. How many species of mammals has Australia lost?

推論質問

Q1. Why does a “shocking” loss of nature happen?

Q2. Why is the report depressing?

Q3. Why is the environment in Australia getting worse and worse?

Q4. Why is the report like a red alert?

Q5. Why does a nature group want to save Australia’s nature?

資料2：足場がけの質問の有無が IU再生率に及ぼす影響について（線形混合モデルを使用）

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	47.58	[41.23, 53.93]	< .001***
With Questions (vs. Without Questions)	5.58	[1.18, 9.99]	.014*
Proficiency	7.00	[1.49, 12.51]	.013*
HATS (vs. ROCKET)	5.50	[0.34, 10.65]	.037*
AUSTRALIA (vs. ROCKET)	-0.79	[-5.93, 4.34]	.759
Random Effects			
σ^2	111.69		
τ00 Participant	140.04		
τ11 Proficiency	8.66		
ρ01 Participant	1.00		
ICC	0.52		
N participant	36		
Observations	103		
Marginal R^2/Conditional R^2	0.190/0.610		

Note. *p < .05, ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients. *Proficiency* represents the reading proficiency test scores (z-scores).

資料3: 週の推移が発話流暢性に与える影響について (線形混合モデルを使用)

(1)総シラブル数

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	93.48	[86.61, 100.35]	< .001***
week 5 (vs. week 1)	6.12	[1.77, 10.47]	.006**
Random Effects			
σ^2	181.43		
τ00 Participant	724.95		
ICC	0.80		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.010/0.802		

Note. **p < .01, ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients.

(2)1秒間の平均シラブル数

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	2.48	[2.40, 2.55]	< .001***
week 5 (vs. week 1)	0.06	[-0.01, 0.13]	.073
Proficiency	0.15	[0.09, 0.21]	< .001***
Random Effects			
σ^2	0.05		
τ00 Participant	0.05		
ICC	0.53		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.213/0.628		

Note. ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients. *Proficiency* represents the reading proficiency test scores (z-scores).

(3)総単語数

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	58.84	[54.57, 63.11]	< .001***
week 5 (vs. week 1)	9.07	[5.88, 12.25]	< .001***
Proficiency	10.91	[7.30, 14.52]	< .001***
Random Effects			
σ^2	97.34		
τ00 Participant	241.20		
ICC	0.71		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.320/0.805		

Note. ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients. *Proficiency* represents the reading proficiency test scores (z-scores).

(4) 1分間の平均単語数

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	96.03	[92.80, 99.26]	< .001***
week 5 (vs. week 1)	11.11	[8.07, 14.14]	< .001***
Proficiency	8.33	[5.74, 10.93]	< .001***
Random Effects			
σ^2	88.40		
τ00 Participant	105.55		
ICC	0.54		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.366/0.711		

Note. ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients. *Proficiency* represents the reading proficiency test scores (z-scores).

(5) ポーズ間の平均シラブル数

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	4.02	[3.72, 4.32]	< .001***
Random Effects			
σ^2	0.69		
τ00 Participant	1.37		
ICC	0.66		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.000/0.665		

Note. ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients.

(6) 平均ポーズ時間

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	1.86	[1.71, 2.02]	< .001***
week 5 (vs. week 1)	-0.29	[-0.46, -0.12]	.001**
Proficiency	-0.16	[-0.28, -0.04]	.009**
Random Effects			
σ^2	0.27		
τ00 Participant	0.17		
ICC	0.39		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.103/0.456		

Note. **p < .01, ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients. *Proficiency* represents the reading proficiency test scores (z-scores).

(7)発話時間の割合

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	45.59	[42.98, 48321]	< .001***
week 5 (vs. week 1)	2.16	[0.01, 4.31]	.051
Proficiency	4.01	[1.83, 6.18]	< .001***
Random Effects			
σ^2	44.44		
τ00 Participant	82.76		
ICC	0.65		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.135/0698		

Note. ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients. *Proficiency* represents the reading proficiency test scores (z-scores).

(8)1分間の平均言いよどみ数

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	7.84	[6.52, 9.17]	< .001***
Random Effects			
σ^2	16.66		
τ00 Participant	25.34		
ICC	0.60		
N participant	75		
Observations	150		
Marginal R^2/Conditional R^2	0.00/0.603		

Note. ***p < .001, CI = confidence interval, ICC = Intraclass correlation coefficients.

資料4：再話の有無が心的表象の強度に及ぼす影響について（線形混合モデルを使用）

	Accuracy		
Predictors	Estimates	95%CI	*p*
(Intercept)	0.86	[0.61, 1.11]	< .001***
Paraphrase (vs. Distractor)	3.79	[3.50, 4.08]	< .001***
Inference (vs. Distractor)	3.04	[2.74, 3.34]	< .001***
With Retelling (vs. Without Retelling)	-0.41	[-0.71, -0.12]	.006**
Proficiency	-0.10	[-0.22, 0.03]	.139
Paraphrase * With Retelling	0.74	[0.39, 1.09]	< .001***
Inference * With Retelling	0.26	[-0.10, 0.62]	.154
Random Effects			
σ^2	2.48		
τ00 Participant	0.04		
τ11 Proficiency	0.12		
ρ01 Participant	0.23		
ICC	0.07		
N participant	104		
Observations	2432		
Marginal R^2/Conditional R^2	0.544/0.575		

Note. $^{**}p < .01$, $^{***}p < .001$, CI = confidence interval, ICC = Intraclass correlation coefficients. *Proficiency* represents the reading proficiency test scores (z-scores).

資料5：足場がけの質問の有無が心的表象の強度に及ぼす影響について（一般化線形混合モデルを使用）

	Accuracy		
Predictors	Odds Ratios	95%CI	*p*
(Intercept)	0.01	[0.00, 0.04]	< .001***
Paraphrase (vs. Distractor)	752.89	[219.80, 2578.84]	< .001***
Inference (vs. Distractor)	138.30	[41.80, 457.55]	< .001***
Random Effects			
σ^2	3.29		
τ00 Participant	0.13		
τ11 Proficiency	0.16		
ρ01 Participant	-0.91		
ICC	0.04		
N participant	36		
Observations	848		
Marginal R^2/Conditional R^2	0.679/0.690		

Note. $^{***}p < .001$, CI = confidence interval, ICC = Intraclass correlation coefficients.

第35回 研究助成
B 実践部門 ● 報告Ⅱ ● 英語能力向上をめざす教育実践

日本人高校生に対する,短期間でより効果的なスピーキング活動:「タスクの繰り返し」の検証

研究者:上野 正和　福岡県／北九州市立高等学校　教諭・茨城県／筑波大学大学院　在籍
《研究助言者:小泉 利恵》

概要　本研究は発話流暢性の向上を調査したものである。日本人高校生を対象とし,絵を見てその内容を英語で描写するタスクの繰り返しを週3回,2週間という比較的短い期間で実施した。そして,どのような発話流暢性の向上が図れるかについて調査した。また,ブロック型(1回目:AAA,2回目:BBB,3回目:CCC)とインターリーブ型(1回目:ABC,2回目:ABC,3回目:ABC)という学習方法を1週目と2週目で様々な組み合わせで行った場合,どのような型の組み合わせで行うことがより発話流暢性の向上に効果的かを調査した。その結果,発話流暢性では「発話の割合」,「文と文の間のポーズの長さ」,「文と文の間のポーズの数」の3観点に向上が見られ,型の組み合わせでは1週目をインターリーブ型で行い,2週目もインターリーブ型で行う「インターリーブ型・インターリーブ型」が最も効果的であることが示唆された。

1 はじめに

日本の高校の教育が今,大きく変わろうとしている。学習指導要領の改訂から5年経つ(文部科学省,2018)が,それに伴い2022年から新教育課程の導入で,3観点による観点別学習状況の評価が厳密に行われるようになった。また大学入試ではセンター試験から2020年度に大学入学共通テストへと代わり,より思考力・判断力等を求められるようになった。その流れの中で,英語教育も大きく変わろうとしている。

伝統的な日本の英語教育は「読むこと」を中心としてきたが,センター試験でのリスニングテスト導入や英検での英作文によるライティングの評価等によって,「聞くこと」や「書くこと」の重要性が示された。そして今,4技能で残された「話すこと」についても,様々な取り組みが行われるようになった。東京都では2022年度より高校入試にスピーキングテストを導入し,評価に「話すこと」を追加した。また,授業等でもインターネットを介したリモートによる国際交流も可能となり,画面を通して外国人と英語を介したコミュニケーションをオンラインで行うことができるようになった。また,これまで主観的になりがちであったために難しいとされてきたスピーキングの評価はAIが一部行うことが可能になった。このように時代の流れと科学技術との融合により,今後は英語教育の中でも,特に「話すこと」の育成が益々重要視されることになるだろうと筆者は考える。

それでは「話すこと」を育成するためにはどのような手立てが必要であろうか。「話すこと」の指針としてCAF(Skehan, 1998)があるが,これはComplexity(複雑性),Accuracy(正確性),Fluency(流暢性)のそれぞれの頭文字をとっている。『高等学校学習指導要領解説 外国語編・英語編』(文部科学省,2018)では,課題としてコミュニケーション能力の向上が挙げられている(p. 7)。ただし,平成29年度英語教育改善の

ための英語力調査の事業報告（文部科学省，2017）の中で，高校3年生の英語力調査結果ではスピーキング能力は大半がCEFRレベルのAであることが分かった（A1が87.2%，A2が11.7%，B1が1.2%）。CEFRレベルがAということは，現状の日本人高校生のスピーキング能力の大半が基礎段階の言語使用者ということになる。別の言い方をすると，大半の高校生は英語を使って自分の伝えたい情報を相手に上手く伝えることができていないとも言える。このような現状を踏まえると，スピーキングにおけるコミュニケーション能力の向上には，込み入った事を伝えることができる複雑性や，間違わずに伝えることができる正確性も大切であると考えられるが，一定時間内に相手にどれだけ多くのことを伝えることができるかがカギの一つとなる「流暢性」の育成が日本人高校生のスピーキング能力の向上には最も必要であると筆者は考える。それでは「流暢性」を向上させるためにはどのような手立てが必要なのか。本調査の主目的である，日本人高校生に対する短期間でより効果的なスピーキング活動の検証について，以下に先行文献の紹介と共に検討していきたい。

2 先行文献

2.1 流暢性について

Skehan（2003）は認知的アプローチとして「流暢性」「正確性」「複雑性」等を挙げ，「流暢性」をタスクのパフォーマンス測定の一つとした。そして「流暢性」にはspeed fluency（話す速度），breakdown fluency（ポーズの長さや頻度），repair fluency（言い直し・繰り返しの頻度）の3つ側面があるとした。Tavakoli and Uchihara（2020）は先行研究（Ahmadian, 2011; Kahng, 2014; Kormos, 2006等）をまとめ，これらの3つには発話プロセスには異なる側面（例：speed fluencyには自動化の程度を反映している等）があると指摘した。日本での研究では，Koizumi and In'nami（2014），Koizumi and Katagiri（2009），小泉・栗﨑（2002），Saito et al.（2018），Suzuki et al（2021）等がスピーキングの評価や向上について流暢性の分析を行い，多くの研究成果を残している。

流暢性の向上を具体的に図る手法としては，いくつか考えられる中の一つとして，タスク・ベースの言語指導（TBLT）を挙げることができる。松村（2017）は著書の中で，TBLTの可能性について，「さまざまな場面で適切に機能することのできる学習者を，活動を通して育成しようとしている」と述べている（p.35）。また，TBLTには様々な種類があるが，流暢性を促すものとして，同じタスクを繰り返す「タスクの繰り返し」（Date, 2017; Fukuta, 2016）がある。そこで，以下にそのことについて書かれた文献をいくつか紹介したい。

2.2 流暢性の向上を図るタスクの繰り返し

森（2016）は「L2学習者が言語を産出するときにワーキングメモリ（WM）容量の欠如が起こると，意味の伝達にその内容を消費し，形式に注意をむけることができなくなる」と述べ，その結果「限られたWM容量の中では，L2学習者は一度きりのタスクでは意味と形式の統合は難しいといえる」（p.153）と，タスクの繰り返しの有効性を説いた。

Suzuki（2021）はタスクの繰り返しをすることによって，流暢性の向上を促せるのではないかと考え，実験協力者を募り調査を行った。具体的にはCEFRレベルA2からB1の大学生を対象とし，3種類の絵を用いたストーリーテリング（絵描写タスク）を3日間行った。そして更にタスクの繰り返しを2つのグループ（ブロック型とインターリーブ型）に分けて行った。ブロック型（B型）とは，同じ日の練習で同じ内容のタスクを繰り返し，次の練習で別の内容を繰り返すものである（1日目：A，A，A，2日目：B，B，B，3日目：C，C，C）。一方インターリーブ型（I型）とは，初日に行った複数の異なった種類のタスクを別の日に繰り返すものである（1日目：A，B，C，2日目：A，B，C，3日目：A，B，C）。そしてSuzuki（2021）は初見の内容で事

前，事後テストを行い，3日間の練習期間と共に流暢性がどのように向上していくのかを調査した。分析は，流暢性の3側面（speed fluency，breakdown fluency，repair fluency）から，更にそれらを9項目の観点に細分化して計測した。その結果，タスクの繰り返しの練習の中ではB型の方がI型よりも9観点中4つの観点（「1秒当たりに産出する音節数」，「文内のポーズの長さ」，「文と文の間のポーズの長さ」，「文と文の間のポーズの数」）で優っていた。また事前・事後テストからspeed fluencyの「1秒当たりに産出する音節数」とbreakdown fluencyの4つの観点全てで，より流暢性が向上することを明らかにした（効果量はspeed fluencyが中程度，breakdown fluencyの4つとも小程度であった）。

日本人高校生を対象とした研究では，前述した森（2016）や千菊（2018）が挙げられる。森（2016）は英検準1級の教材を使い，4枚1組の絵の「絵描写タスク」を3回取り組ませた。そしてタスクの繰り返しを行うことでCAFがどの程度向上したかを分析した。その結果，流暢さについては「発話の速度とポーズ」の合計について有意差が見られた。千菊（2018）は森（2016）と同様にストーリーテリングの課題に取り組ませたが，森（2016）が実験を実施した回数が3回であったのに対し，千菊（2018）は3回のタスクの繰り返しを合計16回実施した。そしてその前後の発話の質の変容をタスクの繰り返しをしていない別クラスと比較した。結果は流暢さでは「1分間あたりの実発話語数」に向上が見られた。

しかし流暢性には前述した通り，speed fluencyとbreakdown fluencyとrepair fluencyの3つの側面があるが，森（2016）はspeed fluencyとbreakdown fluency，千菊（2018）はspeed fluencyのみを測っており，両者とも3側面全ての流暢性を計測していない。一方Suzuki（2021）は流暢性についてspeed fluency，breakdown fluency，repair fluencyの3つの側面を更に詳細に9つの観点に分けて詳しく分析していたが，タスクの繰り返しの回数は3回×3日間しかしていない。また，Suzuki（2021）の対象は高校生ではなく大学生である。他の文献も調べてみたが，高校生の流暢性について詳細に分析したものは見当たらなかった（日本人高校生以外を対象とした例は，De Jong & Perfetti，2011；Vercellotti，2017等）。

以上のことを踏まえ，本研究では高校生を対象に，流暢性を詳細に分析したSuzuki（2021）を基に，先ず3回のタスクの繰り返しを1週間で3回行い，タスクの前後で初見の内容でテスト（事前テスト・中間テスト）を行い，Suzuki（2021）と同様な流暢性の向上が見られるかどうかを調査する。そしてSuzuki（2021）の研究よりも期間を延長し，その後更に1週間（3回）タスクの繰り返しを行い，合計6回タスクの繰り返しを行う。また，6回目のタスクの繰り返しが終了した後に，初見の内容で事後テストを行い，流暢性を測り，事前・中間テストと比較することで流暢性がどのように向上していくかを調査する。タスクのグループ分けの方法もSuzuki（2021）に倣い，B型とI型の2種類の型に分けることにした。ただし，Suzuki（2021）ではB型とI型の2グループであったが，本調査では更に発展させ，3回ずつを1ユニットとし，「ブロック型・ブロック型」（BB型）「インターリーブ型・インターリーブ型」（II型）に加え，「ブロック型・インターリーブ型」（BI型）及び「インターリーブ型・ブロック型」（IB型）の4つのグループに分け，どのグループが最も流暢性が向上するか調査する。このようにしてSuzuki（2021）よりも長く複雑な調査を行うことにより，週3回を2週間という比較的短期間であるが，より詳細な調査を試みることとした。以上のことから，研究課題（RQ）は以下の2点となる。

RQ1　日本人高校生がタスクの繰り返しを行った際に，初見の内容で，流暢性のどの側面が向上するのか。

RQ2　タスクの繰り返しをする場合，どのような型の組み合わせが最も効果的に流暢性を向上させるのか。

3 方法

3.1 協力者

協力者は公立高校に通う3年生119名の生徒だった(3クラスに所属)。この内, 1クラス(39名)をパイロットテスト用のクラス, そして残りの2クラス(80名)を本研究での分析対象とした。パイロットテスト用のクラスでは, どのようなマテリアルでタスクの繰り返しを行うのが適当かを調査した。そして残る2クラスでは, 事前・中間・事後テストと, その間に6回のタスクの繰り返しの練習を行った。分析はタスクの繰り返しの練習やテストにすべて参加した者を対象に行い, 結果的に58名を調査の対象とした。

58名の特徴としては, アンケート等で調査を行った結果, 学力はCEFRレベルでA1～A2であるが, 大半の生徒がA1と推定された。教科として英語を学び始めたのは中学校1年生からであるが, 外国語活動として小学校5年生から始めているので, 7年以上英語に慣れ親しんでいる。また, 長期間外国に滞在した経験を持つものはいなかった。英語に関して好感を持っているか尋ねたところ, 約65%の生徒が「あまり好きではない」もしくは「嫌い」と答えたので, 英語に関して苦手意識のある生徒が多いことが窺えた。4技能について得意なものを尋ねたところ, 最も多かったのが「読むこと」で全体の約49%, 次いで「書くこと」約22%, 「聞くこと」約21%, 「話すこと」約8%の順となった。一方, 4技能でどの技能を習得したいかについては, 最も多かったのが「話すこと」で全体の約55%, 次いで「聞くこと」約23%, 「読むこと」約12%, 「書くこと」約10%の順となった。生徒の最も苦手な技能が「話すこと」であったが, またその反面, 最も習得したい技能が同じ「話すこと」であったことは大変興味深く感じられた。

3.2 マテリアル

調査で使用するマテリアルの内容を決めるために, パイロットテスト用として, 英検3級の2次試験問題の第5問, 英検準2級の2次試験問題の第3問, 第4問, 第5問の4種類の問題を用意した。そして生徒にそれぞれの問題について回答をした時の感想を書いてもらった。当初は, 絵以外の質問形式も検討したので, 英検3級の第5問, 英検準2級の第4, 5問といった質問に対して答える形式の問題も用意した。そして英検準2級の第3問の絵を見てその内容を描写する問題も含め, どの形式が一番良いのか検討した。また, 絵を見てその内容を描写する問題については, 本調査では協力者の英語力を考慮して, 森(2016)やSuzuki(2021)の様に複数の絵を見せて描写するものではなく, 1枚の絵を見て内容を描写する英検準2級のレベルに留めた。

パイロットテストの結果, 英検3級の第5問と英検準2級の第3問が発話しやすかったという生徒が多かった。理由として, 前者は身近で平易な内容であることであり, 後者は問題が絵を見ながら答えられるので英語を発話しやすかったということであった。しかし, 英検3級の内容であると発話内容が極めて限られてしまうため発話時間を長く得にくいという理由で却下した。したがって, 本調査では英検準2級二次試験問題の第3問を使用することを決定した(資料1参照)。

本調査では事前・中間・事後テスト用にそれぞれ1種類ずつ, そしてタスクの繰り返し用に2週間分(3回×2週×2種類)の合計15種類(事前テスト, 中間テスト, 事後テスト, A, A', B, B', C, C', D, D', E, E', F, F')のマテリアルを用意した。タスクの繰り返し用に2種類ずつ(例:A, A')用意した理由は, ペアで練習させるために同じマテリアルであると後半に行う生徒が初見でないために前半の生徒の言ったことを反復してしまう可能性があると考えたからである。また, マテリアルの絵の内容については全て英検準2級の二次問題の第3問から出題したので, 同レベルの問題と解釈し, カウンターバランスは行わなかった。

3.3 手順

手順に関しては表1の流れで行った。事前テストは同僚の英語科の4人の教員に手伝ってもらい，1クラスを5つ(8人程度)に分け，別室へ連れていき個別にICレコーダーで発話内容を録音した。即興での発話の流暢性を調べるために，発話を計画する時間は設けなかった。発話時間は一人1分30秒で，絵の内容を描写させた。タスクの繰り返しの練習の時は1分間で行ったが，テストではなるべく多くの発話を期待して1分半とした。テストを終えた生徒は教室へ戻り，事前アンケートの質問事項に回答した(資料2参照)。

タスクの繰り返しの練習は，授業時間1回50分を使い，週3回・合計6回実施した(表2参照)。対象の2クラス(クラスAとB)について，1週目(1〜3回)には，各クラスをそれぞれB型とI型に20名ずつ分けた。2週目(4〜6回)には，クラスAでは1週目(1〜3回)と2週目(4〜6回)を異なる型で行い，クラスBでは1週目と2週目を同じ型で行った(表3参照)。

■表1: 本調査の流れ

・事前テスト → ・事前アンケート → ・タスクの繰り返し1回目 → ・タスクの繰り返し2回目 → ・タスクの繰り返し3回目 → ・中間テスト → ・タスクの繰り返し4回目 → ・タスクの繰り返し5回目 → ・タスクの繰り返し6回目 → ・事後テスト → ・事後アンケート

■表2: 各型の使用マテリアルとマテリアルの使用順序

	1回目	2回目	3回目	4回目	5回目	6回目
ブロック型	AAA A' A' A'	BBB B' B' B'	CCC C' C' C'	DDD D' D' D'	EEE E' E' E'	FFF F' F' F'
インターリーブ型	ABC A' B' C'	ABC A' B' C'	ABC A' B' C'	DEF D' E' F'	DEF D' E' F'	DEF D' E' F'

■表3: クラス別ブロック型とインターリーブ型の1週目と2週目

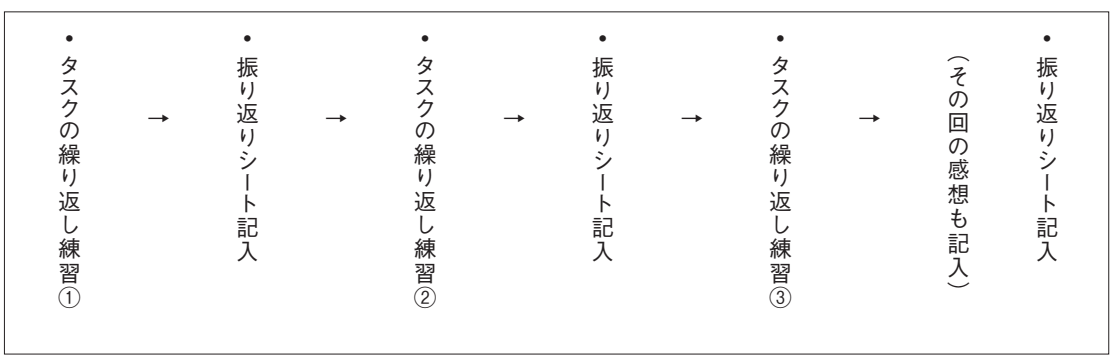
	1週目(1 〜3回)	2週目(4 〜6回)
クラスA	ブロック型	インターリーブ型
	インターリーブ型	ブロック型
クラスB	ブロック型	ブロック型
	インターリーブ型	インターリーブ型

タスクの繰り返しの練習は表4の流れで行った。また，練習中はペアで行った。B型とI型でそれぞれ定期考査等の成績を参考に学力順に並べ，英語が得意な生徒と不得意な生徒でペアを組ませることにした。理由は異なった学力差で組ませた方が，よりピアサポートが図れ，より流暢性の向上に繋がるのではないかと考えたからである。また，練習時には黒板に向かって左側に英語が得意な生徒，そして右側に英語が得意ではない生徒が座るように配置し，タスクの繰り返し練習①を終えた後に，英語が得意な生徒の列の席

を一つ前にずらすことで異なるペアを組めるようにし，ペア同士が固定しないようにした。欠席者がB型もしくはI型内に偶数出た場合は，ペアがいない者同士で組ませ，奇数の場合はT2（チーム・ティーチング授業における2名の先生のうち，補助の先生）がペアになった。具体的な練習方法として，先ず練習開始時にじゃんけんをして，勝った方から1分間渡されたマテリアルの絵について英語で発話した。こうすることにより，教室内で一斉に行っていても，隣が同じものを行うとは限らないので，真似をすることを防ぐことができた。また，発話するときには発話者はペアを意識して，相手にしっかり伝えるように指導した。そして，発話をしていない方には，うなずく等の介入が入ると純粋な発話実験にならないので，黙って真剣に相手の話を聴くように指導した。1分が経ち発話が終わると役割を交代させた。また，生徒はより自然な発話ができるように，マテリアルで1度話した後には（例：A，A’の順に話した後に）模範解答（英検の過去問題集，2012, 2016, 2017, 2018の解答から抜粋したもの）を渡して参考とした。

タスクを行った後に毎回5分間，タスクの内容の復習と，どのようなことを学んだのかを振り返るために「振り返りシート」に記入した（資料3参照）。そしてタスクの繰り返し練習③まで終えた後にその回の振り返りとして，学んだことや感想を書いた。なお，上述した模範解答については，この5分間の中で配布し内容確認することとした。

■表4: 毎回のタスクの繰り返しの流れ



（注）各「振り返りシート記入」時には，マテリアル初見の場合のみ模範解答を配布した。

生徒の多くがスピーキングを苦手としていることが事前アンケートの結果から分かったので，毎回のタスク活動後にスピーキングの練習を行った。ただし，4回目と6回目の後にはそれぞれ中間テストと事後テストを行ったので，その2回はスピーキング練習をすることはできなかった（表5参照）。また，事後テスト後に本調査について事後アンケートを行った（資料4参照）。

■表5: タスク活動後に行ったこと

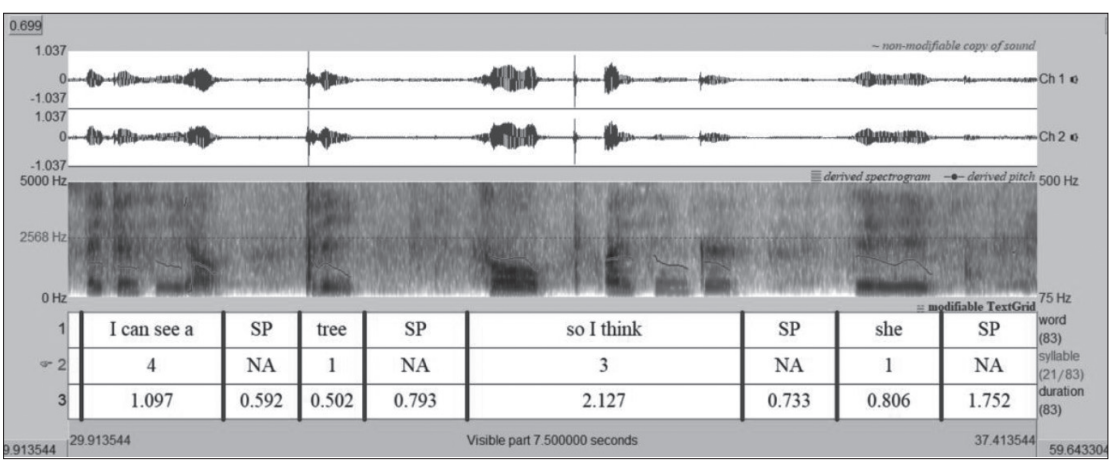
回数	内容
1回目	①インターネットを使った辞書の引き方，翻訳機の使い方について （※ただし翻訳機は万能ではないことも留意させる）の説明 ②表現の種類（感情・描写）の説明 ③状況の描写について具体的な説明（いつ・どこで・だれ・何をする） ④日本語と英語の語順の違いについて
2回目	「だれ（何）」についての英語表現の練習（目に入ったものに対して言うことができる表現） （”I can see ～ in the picture.”）
3回目	中間テスト

回数	内容
4回目	「その人が…にいると思う」についての英語表現 の練習（場所と…と思うについての表現） （“So I think he/she is in ～ .”）
5回目	「その人は～したいけど…」についての英語表現練習（願望と接続詞についての表現） （He /She wants to… but….）
6回目	事後テスト，事後アンケート

3.4 分析

事前・中間・事後テストにおいてICレコーダーで録音した発話ファイルをPRAAT（Boersma & Weenink, 2020）で分析した（図1参照）。PRAATとは，音声を分析，変換，合成することができるフリーソフトウェアで，オランダのアムステルダム大学のPaul Boersma氏とDavid Weenink氏によって開発されたものである。また，無償で手に入れることができ，比較的短時間で容易に再合成音声や実験が準備できる（北原・田嶋，2011; 名倉，2010）。分析は筆者一人で行ったが，分析時間に約3か月を費やし，長い時間をかけることによって，丁寧に分析することを心がけ，事前・中間・事後テストの分析の仕方に隔たりがなく整合性があるかを意識し，3回以上見直しをした。分析観点はSuzuki（2021）に倣い，流暢性の3側面を9つの観点に分けた（表6, 7参照）。接続詞を伴う文はThe Analysis of Speech Unit（AS-unit; Foster et al., 2000）を採用し，音調の後に0.5秒以上のポーズがある場合には別のAS-unitとして区切った。書き起こしは，最後に長いポーズがあった場合でも90秒間書き起こした。

PRAATから算出した数値に対して，SPSSで統計分析を行った。2要因被験者間・被験者内混合分散分析を用いて，各テスト間（事前・中間・事後テスト）と各型（B型，I型等）を分析した。また，Suzuki（2021）と比較するために，事前・中間テスト間（タスクの繰り返しの1回目から3回目の前後）で分析した後に，事前・中間・事後テスト間（タスクの繰り返しの1回目から3回目の前後と6回目の後）を分析した。効果量はSuzuki（2021）を参考に，偏イータ2乗（Richardson, 2011）を採用し，有意確率はp<.05とした。事前・中間テスト間ではサンプルサイズが30程度であったため，球面性検定ではグリーンハウス・ガイザー補正を採用したが，事前・中間・事後テスト間ではサンプルサイズが10から17と小さかったので，ホイン・フェルト補正を採用した。外れ値もSuzuki（2021）に倣い，z > 3.29（Tabachnick & Fidell, 2013）を基準とし分析を行った。その結果，観点1（音節数の平均）の中で3名の外れ値があったので分析から外した。



■図1：PRAATを用いた音声編集の一例

■表6: 流暢性の9観点

Speed fluency	1. (ポーズとポーズの間の) 音節数の平均 (単位:音節) (mean length of run)	ポーズとポーズの間の音節数(発話中の不完全な単語も含む一方,「ええと…」などの表現は除外)÷ポーズとポーズの間の固まり数
	2. 1秒当たりに産出する音節数 (単位:音節) (articulation rate)	90秒間で発話した音節数 ÷ 発話のトータルの長さ(ポーズを除いた発話時間)
	3. 発話の割合 (単位:%) (phonation/time ratio)	発話のトータルの長さ(ポーズを除いた発話時間) ÷ 全体の長さ(90秒)×100
Breakdown fluency	4. 文内のポーズの長さ (単位:秒) (mid-clause pause duration)	発話中のAS-unit内の単語と単語の間でポーズがあったときのトータルのポーズの長さ
	5. 文と文の間のポーズの長さ(単位:秒) (clause-final pause duration)	発話中のAS-unitとAS-unitの間でポーズがあったときのトータルのポーズの長さ
	6. (1分当たりの) 文内のポーズの数(単位:回) (mid-clause pause frequency)	発話中のAS-unit内の単語と単語の間でポーズがあったときのトータルのポーズ数 ÷ 全体の長さ(90秒)×1分(60秒)
	7. (1分当たりの) 文と文の間のポーズの数 (単位:回) (clause-final pause frequency)	発話中のAS-unitとAS-unitの間でポーズがあったときのトータルのポーズ数 ÷ 全体の長さ(90秒)×1分(60秒)
Repair fluency	8. (1分当たりの) 繰り返しの数(単位:回) (repetition frequency)	繰り返しの数 ÷ 全体の長さ(90秒)×1分(60秒)
	9. (1分当たりの) 言い直しの数(単位:回) (repair fluency frequency)	言い直しの数 ÷ 全体の長さ(90秒)×1分(60秒)

■表7: 発話分析の一例(図1の内容を9観点で表した場合)

発話内容	I can see a … tree. … So, I think … she …[続く].
観点1 「(ポーズとポーズの間の) 音節数の平均」	"I can see a" が4音節, "tree" が1音節, "So I think" が3音節, "she" が1音節となり, ポーズとポーズの間の固まり数が4つとなるので, 観点1は(4+1+3+1)÷4=2.25音節となる。
観点2 「1秒当たりに産出する音節数」	総音節数が(4+1+3+1)=9となり, 発話のトータルの長さが(I can see a (1.097[秒])+tree(0.502)+So I think(2.127)+she(0.806))=4.532となるので, 観点2は9÷4.532=1.986音節となる。
観点3 「発話の割合」	発話のトータルの長さが4.532となり, 図1の全体の長さが1.097+0.592+0.502+0.793+2.127+0.733+0.806+1.752=8.402となるので, 観点3は4.532÷8.401×100=53.946%となる。
観点4 「文内のポーズの長さ」	I can see aとtreeの間のポーズの長さが0.592であり, So, I thinkとsheと次の単語の間のポーズの長さが0.733と1.752なので, 観点4は0.592+0.733+1.752=3.077秒となる。
観点5 「文と文の間のポーズの長さ」	I can see a tree.とSo, I think she….の間のポーズの長さが0.793なので, 観点5は0.793秒となる。
観点6 「(1分当たりの) 文内のポーズの数」	トータルのポーズの数(I can see aとtreeの間に1つ, So, I thinkとsheとその後の単語の間に2つ)=3つになり, 全体の長さが8.401となるので, 観点6は3÷8.401×60=21.426回となる。
観点7 「(1分当たりの) 文と文の間のポーズの数」	トータルのポーズの数(I can see a tree.とSo, I think she….との間に1つ)=1つになり, 全体の長さが8.401となるので, 観点7は1÷8.401×60=7.142回となる。
観点8「(1分当たりの) 繰り返しの数」	図2の中に繰り返しの表現がないので, 観点8は0÷8.401×60=0回となる。
観点9「(1分当たりの) 言い直しの数」	図2の中に言い直しの表現がないので, 観点9は0÷8.401×60=0回となる。

(注) I can I can can see … と言った場合には, 繰り返し数は2回, 言い直しは1回とした。

4 結果

流暢性の9つの観点に関して事前・中間・事後テストの平均値及び標準偏差を産出した（表8参照）。

■表8: 記述統計量

	事前テスト			中間テスト			事後テスト		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
1. 音節数の平均	55	2.198音節	0.615	55	2.325音節	0.472	55	2.420音節	0.531
2. 1秒当たりに産出する音節数	58	2.150音節	0.539	58	1.972音節	0.360	58	2.066音節	0.352
3. 発話の割合	58	11.640%	0.065	58	23.867%	0.082	58	25.862%	0.075
4. 文内のポーズの長さ	58	21.580秒	17.261	58	28.447秒	12.469	58	30.520秒	12.785
5. 文と文の間のポーズの長さ	58	57.695秒	19.229	58	40.081秒	15.516	58	35.814秒	15.723
6. 文内のポーズの数	58	4.989回	3.216	58	8.230回	3.739	58	9.264回	3.700
7. 文と文の間のポーズの数	58	2.874回	1.097	58	4.805回	1.613	58	5.207回	1.288
8. 繰り返しの数	58	0.724回	1.047	58	1.667回	1.605	58	1.805回	1.766
9. 言い直しの数	58	0.506回	0.731	58	0.828回	0.859	58	0.989回	1.090

（注）音節数の平均のみ外れ値が，事前・中間，事前・事後，中間・事後テスト間において，それぞれ1名分あったので，合計3人を除外し，n=55となっている。

1週目（1回～3回）の前後で行った事前テストと中間テストで，流暢性を測る9観点を分析した結果（表9左側「事前・中間テスト」の分散分析結果参照），テスト間の要因は9つの観点の内，観点2から観点9が有意であったのに対し，型（B型・I型）間の有意確率は観点1以外は有意ではなかった。また，テストと型の交互作用での有意確率も全ての観点で有意ではなかった。効果量については上述したようにSuzuki（2021）に倣い，偏イータ2乗（Richardson, 2011）を採用し，効果量を小: $\eta p^2 = .0099$; 中: $\eta p^2 = .0588$; 大: $\eta p^2 = .1379$とした。その結果，テスト（事前・中間）間に対して流暢性の効果量が大きかったものが5つ（観点3, 5, 6, 7, 8），中が3つ（観点2, 4, 9）小が1つ（観点1），そして効果量無しが0であった。一方，型（B型・I型）間の効果量は大が0，中が1つ（観点1），小が4つ（観点4, 6, 8, 9）で，効果量無しが4つ（観点2, 3, 5, 7）であった。また，テストと型の交互作用に対しては大が0，中が0，小が5つ（観点2, 3, 6, 7, 9）で，効果量無しが4つ（観点1, 4, 5, 8）であった。

1週目の後に2週目（4回～6回）を行い，事後テストを含めて分析した結果（表9右側「事前・中間・事後テスト」の分散分析結果参照），事前・中間・事後のテスト間の要因は9つの観点の全ての観点で有意だったのに対し，型（BB型・II型・BI型・IB型）間の有意水準は9の観点全てで有意ではなかった。また，テストと型の交互作用の有意確率も全ての観点で有意ではなかった。効果量については，テスト（事前・中間・事後）間に対して大が5つ（観点3, 5, 6, 7, 8），中が4つ（観点1, 2, 4, 9），小が0，効果量無しが0であった。一方，型（BB型・II型・BI型・IB型）間の効果量は大が0，中が1つ（観点1），小が8（観点2～10）で，効果量無しが0であった。テストと型の交互作用に対しては大が0，中が5つ（観点1, 2, 3, 6, 7, ），小が4つ（観点4, 5, 8, 9）で，効果量無しが0であった。

次にほとんどの観点で有意差があったテスト間について，流暢性の9つの観点ごとに，事前・中間テストと事前・中間・事後テストの2つにテストに分けて，各型について単純主効果の分析を，多変量分散分析を用いて詳細に行った（表10～18参照）。また，サンプルサイズが比較的小さかったため，Pillaiのトレースの結果を採用した。

1つ目の観点である「音節数の平均」の多変量検定を行った結果，事前・中間テスト間では，B型（$F(1, 53) = 1.401$, $p = .242$, $\eta p^2 = .026$）；I型（$F(1, 53) = 0.699$, $p = .407$, $\eta p^2 = .013$）ともに有意ではなく，効果量も小さかった（表10参照）。

事前・中間・事後テスト間では，BB型（$F(2, 50) = 1.712$, $p = .191$, $\eta p^2 = .064$）；II型（$F(2, 50) = 8.958$, $p < .001$, $\eta p^2 = .264$）；BI型（$F(2, 50) = 1.611$, $p = .210$, $\eta p^2 = .061$）；IB型（$F(2, 50) = 2.330$, $p = .108$, $\eta p^2 = .085$）となり，II型のみが有意であった。また，効果量はII型が大となり，その他の型は中程度となった（表10参照）。

■表9：事前・中間テストと事前・中間・事後テストの2元配置分散分析

事前・中間テスト

1．音節数の平均

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	1.987	1	1.987	5.434	0.024	0.093（中）
誤差	19.376	53	0.366			
Within Subjects						
テスト	0.455	1	0.455	2.067	0.156	0.038（小）
テスト・型	0.021	1	0.021	0.097	0.757	0.002（無）
誤差(テスト)	11.662	53	0.22			
合計	33.501	109				

2．1秒当たりに産出する音節数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	9.96E−08	1	9.96E−08	0	1	0（無）
誤差	16.241	56	0.29			
Within Subjects						
テスト	0.891	1	0.891	6.627	0.013	0.106（中）
テスト・型	0.168	1	0.168	1.251	0.268	0.022（小）
誤差(テスト)	7.526	56	0.134			
合計	2.48E+01	115				

3．発話の割合

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	0	1	0	0.027	0.871	0（無）
誤差	0.472	56	0.008			
Within Subjects						
テスト	0.43	1	0.43	170.204	<.001	0.752（大）
テスト・型	0.004	1	0.004	1.414	0.239	0.025（小）
誤差(テスト)	0.142	56	0.003			
合計	1.048	115				

4．文内のポーズの長さ

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	159.53	1	159.53	0.609	0.438	0.011（小）
誤差	14666.95	56	261.91			
Within Subjects						
テスト	1285.392	1	1285.392	6.722	0.012	0.107（中）
テスト・型	8.968	1	8.968	0.047	0.829	0.001（無）
誤差(テスト)	10707.62	56	191.207			
合計	26828.45	115				

5．文と文の間のポーズの長さ

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	188.533	1	188.533	0.472	0.495	0.008（無）
誤差	22349.09	56	399.091			
Within Subjects						
テスト	8917.467	1	8917.467	41.101	<.001	0.423（大）
テスト・型	111.985	1	111.985	0.516	0.475	0.009（無）
誤差(テスト)	12150.1	56	216.966			
合計	43717.17	115				

6．文内のポーズの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	18.983	1	18.983	1.058	0.308	0.019（小）
誤差	1005.109	56	17.948			
Within Subjects						
テスト	299.251	1	299.251	48.639	<.001	0.465（大）
テスト・型	17.886	1	17.886	2.907	0.094	0.049（小）
誤差(テスト)	344.539	56	6.152			
合計	1685.768	115				

7．文と文の間のポーズの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	1.104	1	1.104	0.388	0.536	0.007（無）
誤差	159.441	56	2.847			
Within Subjects						
テスト	106.851	1	106.851	111.486	<.001	0.666（大）
テスト・型	2.634	1	2.634	2.748	0.103	0.047（小）
誤差(テスト)	53.672	56	0.958			
合計	323.702	115				

8．繰り返しの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	1.409	1	1.409	0.559	0.458	0.01（小）
誤差	141.26	56	2.522			
Within Subjects						
テスト	25.578	1	25.578	21.546	<.001	0.278（大）
テスト・型	0.197	1	0.197	0.166	0.685	0.003（無）
誤差(テスト)	66.479	56	1.187			
合計	234.923	115				

9．言い直しの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	2.594	1	2.594	3.168	0.081	0.054（小）
誤差	45.847	56	0.819			
Within Subjects						
テスト	3.082	1	3.082	7.301	0.009	0.115（中）
テスト・型	0.462	1	0.462	1.093	0.3	0.019（小）
誤差(テスト)	23.641	56	0.422			
合計	75.626	115				

（注）偏イータ2乗内の（ ）の文字は効果量の大きさの程度の解釈を表す。

事前・中間・事後テスト

1．音節数の平均

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	2.469	3	0.823	1.583	0.205	0.085（中）
誤差	26.519	51	0.52			
Within Subjects						
テスト	1.166	1.892	0.616	3.459	0.038	0.064（中）
テスト・型	2.365	5.677	0.417	2.338	0.04	0.121（中）
誤差(テスト)	17.19	96.504	0.178			
合計	49.709	158.073				

2．1秒当たりに産出する音節数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	0.325	3	0.108	0.307	0.82	0.017（小）
誤差	19.058	54	0.353			
Within Subjects						
テスト	0.928	1.602	0.579	4.725	0.017	0.08（中）
テスト・型	1.024	4.805	0.213	1.737	0.137	0.088（中）
誤差(テスト)	10.607	86.491	0.123			
合計	31.942	149.898				

3．発話の割合

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	0.018	3	0.006	0.507	0.679	0.027（小）
誤差	0.643	54	0.012			
Within Subjects						
テスト	0.63	2	0.315	137.035	<.001	0.717（大）
テスト・型	0.025	6	0.004	1.808	0.104	0.091（中）
誤差(テスト)	0.248	108	0.002			
合計	1.564	173				

4．文内のポーズの長さ

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	373.722	3	124.574	0.393	0.759	0.021（小）
誤差	17126.59	54	317.159			
Within Subjects						
テスト	2556.506	2	1278.253	8.029	<.001	0.129（中）
テスト・型	505.408	6	84.235	0.529	0.785	0.029（小）
誤差(テスト)	17193.5	108	159.199			
合計	37755.73	173				

5．文と文の間のポーズの長さ

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	347.07	3	115.69	0.232	0.874	0.013（小）
誤差	26957.34	54	499.21			
Within Subjects						
テスト	14482.05	2	7241.023	38.046	<.001	0.413（大）
テスト・型	1030.479	6	171.746	0.902	0.496	0.048（小）
誤差(テスト)	20554.98	108	190.324			
合計	63371.91	173				

6．文内のポーズの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	56.5	3	18.833	0.74	0.533	0.039（小）
誤差	1373.969	54	25.444			
Within Subjects						
テスト	521.241	2	260.62	41.474	<.001	0.434（大）
テスト・型	57.558	6	9.593	1.527	0.176	0.078（中）
誤差(テスト)	678.668	108	6.284			
合計	2687.936	173				

7．文と文の間のポーズの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	5.828	3	1.943	0.57	0.637	0.031（小）
誤差	184.205	54	3.411			
Within Subjects						
テスト	164.418	1.984	82.873	79.503	<.001	0.596（大）
テスト・型	9.766	5.952	1.641	1.574	0.162	0.08（中）
誤差(テスト)	111.676	107.135	1.042			
合計	475.893	172.071				

8．繰り返しの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	11.581	3	3.86	0.945	0.426	0.05（小）
誤差	220.7	54	4.087			
Within Subjects						
テスト	34.27	2	17.135	12.544	<.001	0.189（大）
テスト・型	7.315	6	1.219	0.892	0.503	0.047（小）
誤差(テスト)	147.528	108	1.366			
合計	421.394	173				

9．言い直しの数

Score	SS	df	MS	F	p	ηp2
Between Subjects						
型	1.983	3	0.661	0.575	0.634	0.031（小）
誤差	62.082	54	1.15			
Within Subjects						
テスト	6.187	1.913	3.234	4.649	0.013	0.079（中）
テスト・型	4.375	5.739	0.762	1.096	0.369	0.057（小）
誤差(テスト)	71.87	103.298	0.696			
合計	146.497	167.95				

（注）偏イータ2乗内の（ ）の文字は効果量の大きさの程度の解釈を表す。

■表10：観点1「音節数の平均」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	25	0.026	1.401	1	53	0.242	0.026(小)
I 型	30	0.013	0.699	1	53	0.407	0.013(小)

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	10	0.064	1.712	2	50	0.191	0.064(中)
I I 型	17	0.264	8.958	2	50	<.001	0.264(大)
B I 型	15	0.061	1.611	2	50	0.210	0.061(中)
I B 型	13	0.085	2.330	2	50	0.108	0.085(中)

(注)偏イータ2乗内の( )の文字は効果量の大きさの程度の解釈を表す。

2つ目の観点である「1秒あたりに産出する音節数」の多変量検定を行った結果，事前・中間テスト間では，B型($F(1, 56) = 1.024$, $p = .316$, $\eta p^2 = .018$)；I型($F(1, 56) = 7.062$, $p = .010$, $\eta p^2 = .112$)となり，I型のみ有意であった。効果量はB型が小程度，I型が中程度となった(表11参照)。

事前・中間・事後テスト間では，BB型($F(2, 53) = 2.927$, $p = .062$, $\eta p^2 = .099$)；II型($F(2, 53) = 5.380$, $p = .007$, $\eta p^2 = .169$)；BI型($F(2, 53) = 1.359$, $p = .266$, $\eta p^2 = .049$)；IB型($F(2, 53) = 1.346$, $p = .269$, $\eta p^2 = .048$)となり，II型のみが有意であった。また，効果量はII型が大となり，BB型が中程度，BI型とIB型は小程度となった(表11参照)。

■表11：観点2「1秒当たりに産出する音節数」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.018	1.024	1	56	0.316	0.018(小)
I 型	30	0.112	7.062	1	56	0.010	0.112(中)

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.099	2.927	2	53	0.062	0.099(中)
I I 型	17	0.169	5.380	2	53	0.007	0.169(大)
B I 型	17	0.049	1.359	2	53	0.266	0.049(小)
I B 型	13	0.048	1.346	2	53	0.269	0.048(小)

(注)偏イータ2乗内の( )の文字は効果量の大きさの程度の解釈を表す。

3つ目の観点である「発話の割合」の多変量検定を行った結果，事前・中間テスト間では，B型($F(1, 56) = 67.951$, $p < .001$, $\eta p^2 = .548$)；I型($F(1, 56) = 104.944$, $p < .001$, $\eta p^2 = .652$)ともに有意であった。効果量はB型，I型ともに大であった(表12参照)。

事前・中間・事後テスト間では，BB型($F(2, 53) = 15.037$, $p < .001$, $\eta p^2 = .362$)；II型($F(2, 53) = 55.821$, $p < .001$, $\eta p^2 = .678$)；BI型($F(2, 53) = 36.111$, $p < .001$, $\eta p^2 = .577$)；IB型($F(2, 53) = 19.409$, $p < .001$, $\eta p^2 = .423$)となり，全ての型が有意であった。また，効果量も全ての型で大となった(表12参照)。

■表12：観点3「発話の割合」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.548	67.951	1	56	<.001	0.548（大）
I 型	30	0.652	104.944	1	56	<.001	0.652（大）

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.362	15.037	2	53	<.001	0.362（大）
I I 型	17	0.678	55.821	2	53	<.001	0.678（大）
B I 型	17	0.577	36.111	2	53	<.001	0.577（大）
I B 型	13	0.423	19.409	2	53	<.001	0.423（大）

（注）偏イータ2乗内の（ ）の文字は効果量の大きさの程度の解釈を表す。

4つ目の観点である「文内のポーズの長さ」の多変量検定を行った結果，事前・中間テスト間では，B型（$F(1, 56) = 2.729$, $p = .104$, $\eta p^2 = .046$）；I型（$F(1, 56) = 4.087$, $p = .048$, $\eta p^2 = .068$）ともに有意でなかった。効果量はB型が小程度，I型が中程度となった（表13参照）。

事前・中間・事後テスト間では，BB型（$F(2, 52) = 1.888$, $p = .162$, $\eta p^2 = .068$）；II型（$F(2, 52) = 2.931$, $p = .062$, $\eta p^2 = .101$）；BI型（$F(2, 52) = 1.795$, $p = .176$, $\eta p^2 = .065$）；IB型（$F(2, 52) = 1.099$, $p = .341$, $\eta p^2 = .041$）となり，全ての型で有意でなかった。また，効果量はBB型，II型，BI型が中程度となり，IB型が小程度となった（表13参照）。

■表13：観点4「文内のポーズの長さ」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.046	2.729	1	56	0.104	0.046（小）
I 型	30	0.068	4.087	1	56	0.048	0.068（中）

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.068	1.888	2	52	0.162	0.068（中）
I I 型	17	0.101	2.931	2	52	0.062	0.101（中）
B I 型	17	0.065	1.795	2	52	0.176	0.065（中）
I B 型	13	0.041	1.099	2	52	0.341	0.041（小）

（注）偏イータ2乗内の（ ）の文字は効果量の大きさの程度の解釈を表す。

5つ目の観点である「文と文の間のポーズの長さ」の多変量検定を行った結果，事前・中間テスト間では，B型（$F(1, 56) = 15.663$, $p < .001$, $\eta p^2 = .219$）；I型（$F(1, 56) = 26.322$, $p < .001$, $\eta p^2 = .320$）ともに有意であった。効果量はB型，I型ともに大であった（表14参照）。

事前・中間・事後テスト間では，BB型（$F(2, 53) = 5.666$, $p = .006$, $\eta p^2 = .176$）；II型（$F(2, 53) = 16.673$, $p < .001$, $\eta p^2 = .386$）；BI型（$F(2, 53) = 7.642$, $p = .001$, $\eta p^2 = .224$）；IB型（$F(2, 53) = 3.690$, $p = .032$, $\eta p^2 = .122$）となり，II型のみ有意であった。また，効果量はBB型，II型，BI型が大となり，IB型は中程度であった（表14参照）。

■表14：観点5「文と文の間のポーズの長さ」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.219	15.663	1	56	<.001	0.219（大）
I 型	30	0.320	26.322	1	56	<.001	0.320（大）

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.176	5.666	2	53	0.006	0.176（大）
I I 型	17	0.386	16.673	2	53	<.001	0.386（大）
B I 型	17	0.224	7.642	2	53	0.001	0.224（大）
I B 型	13	0.122	3.690	2	53	0.032	0.122（中）

（注）偏イータ2乗内の（ ）の文字は効果量の大きさの程度の解釈を表す。

6つ目の観点である「文内のポーズの数」の多変量検定を行った結果，事前・中間テスト間では，B 型（$F(1, 56) = 13.419$, $p < .001$, $\eta p^2 = .193$）；I 型（$F(1, 56) = 39.009$, $p < .001$, $\eta p^2 = .411$）ともに有意であった。効果量は B 型，I 型ともに大であった（表15参照）。

事前・中間・事後テスト間では，BB 型（$F(2, 53) = 3.840$, $p = .028$, $\eta p^2 = .127$）；II 型（$F(2, 53) = 22.514$, $p < .001$, $\eta p^2 = .459$）；BI 型（$F(2, 53) = 10.488$, $p < .001$, $\eta p^2 = .284$）；IB 型（$F(2, 53) = 6.855$, $p = .002$, $\eta p^2 = .206$）となり，II 型と BI 型が有意であった。また，効果量は II 型，BI 型，IB 型が大となり，BB 型は中程度であった（表15参照）。

■表15：観点6「文内のポーズの数」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.193	13.419	1	56	<.001	0.193（大）
I 型	30	0.411	39.009	1	56	<.001	0.411（大）

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.127	3.840	2	53	0.028	0.127（中）
I I 型	17	0.459	22.514	2	53	<.001	0.459（大）
B I 型	17	0.284	10.488	2	53	<.001	0.284（大）
I B 型	13	0.206	6.855	2	53	0.002	0.206（大）

（注）偏イータ2乗内の（ ）の文字は効果量の大きさの程度の解釈を表す。

7つ目の観点である「文と文の間のポーズの数」の多変量検定を行った結果，事前・中間テスト間では，B 型（$F(1, 56) = 38.293$, $p < .001$, $\eta p^2 = .406$）；I 型（$F(1, 56) = 77.288$, $p < .001$, $\eta p^2 = .580$）ともに有意であった。効果量は B 型，I 型ともに大であった（表16参照）。

事前・中間・事後テスト間では，BB 型（$F(2, 53) = 13.467$, $p < .001$, $\eta p^2 = .337$）；II 型（$F(2, 53) = 51.769$, $p < .001$, $\eta p^2 = .661$）；BI 型（$F(2, 53) = 42.179$, $p < .001$, $\eta p^2 = .614$）；IB 型（$F(2, 53) = 22.756$, $p < .001$, $\eta p^2 = .462$）となり，全ての型が有意であった。また，効果量も全ての型が大であった（表16参照）。

■表16：観点7「文と文の間のポーズの数」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.406	38.293	1	56	<.001	0.406（大）
I 型	30	0.580	77.286	1	56	<.001	0.580（大）

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.337	13.467	2	53	<.001	0.337（大）
I I 型	17	0.661	51.769	2	53	<.001	0.661（大）
B I 型	17	0.614	42.179	2	53	<.001	0.614（大）
I B 型	13	0.462	22.756	2	53	<.001	0.462（大）

（注）偏イータ2乗内の（　）の文字は効果量の大きさの程度の解釈を表す。

8つ目の観点である「繰り返しの数」の多変量検定を行った結果，事前・中間テスト間では，B型（$F(1, 56) = 8.667$, $p = .005$, $\eta p^2 = .134$）；I型（$F(1, 56) = 13.202$, $p < .001$, $\eta p^2 = .191$）となり，I型のみ有意であった。効果量はB型が中程度，I型が大であった（表17参照）。

事前・中間・事後テスト間では，BB型（$F(2, 53) = 0.658$, $p = .522$, $\eta p^2 = .024$）；II型（$F(2, 53) = 5.469$, $p = .007$, $\eta p^2 = .171$）；BI型（$F(2, 53) = 9.478$, $p < .001$, $\eta p^2 = .263$）；IB型（$F(2, 53) = 3.473$, $p = .038$, $\eta p^2 = .116$）となり，II型，BI型，IB型が有意であった。また，効果量はII型，BI型が大で，IB型が中程度，BB型が小程度であった（表17参照）。

■表17：観点8「繰り返しの数」の多変量検定

事前・中間テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.134	8.667	1	56	0.005	0.134（中）
I 型	30	0.191	13.202	1	56	<.001	0.191（大）

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.024	0.658	2	53	0.522	0.024（小）
I I 型	17	0.171	5.469	2	53	0.007	0.171（大）
B I 型	17	0.263	9.478	2	53	<.001	0.263（大）
I B 型	13	0.116	3.473	2	53	0.038	0.116（中）

（注）偏イータ2乗内の（　）の文字は効果量の大きさの程度の解釈を表す。

9つ目の観点である「言い直しの数」の多変量検定を行った結果，事前・中間テスト間では，B型（$F(1, 56) = 6.788$, $p = .012$, $\eta p^2 = .108$）；I型（$F(1, 56) = 1.421$, $p = .238$, $\eta p^2 = .025$）となり，B型のみ有意であった。効果量はB型が中程度，I型が小程度であった（表18参照）。

事前・中間・事後テスト間では，BB型（$F(2, 53) = 1.555$, $p = .221$, $\eta p^2 = .055$）；II型（$F(2, 53) = 1.471$, $p = .239$, $\eta p^2 = .053$）；BI型（$F(2, 53) = 3.305$, $p = .044$, $\eta p^2 = .111$）；IB型（$F(2, 53) = 2.708$, $p = .076$, $\eta p^2 = .093$）となり，BI型のみが有意であった。また，効果量はBI型，IB型が中程度で，BB型，II型が小程度であった（表18参照）。

■表18: 観点9「言い直しの数」の多変量検定

事前・中間テスト

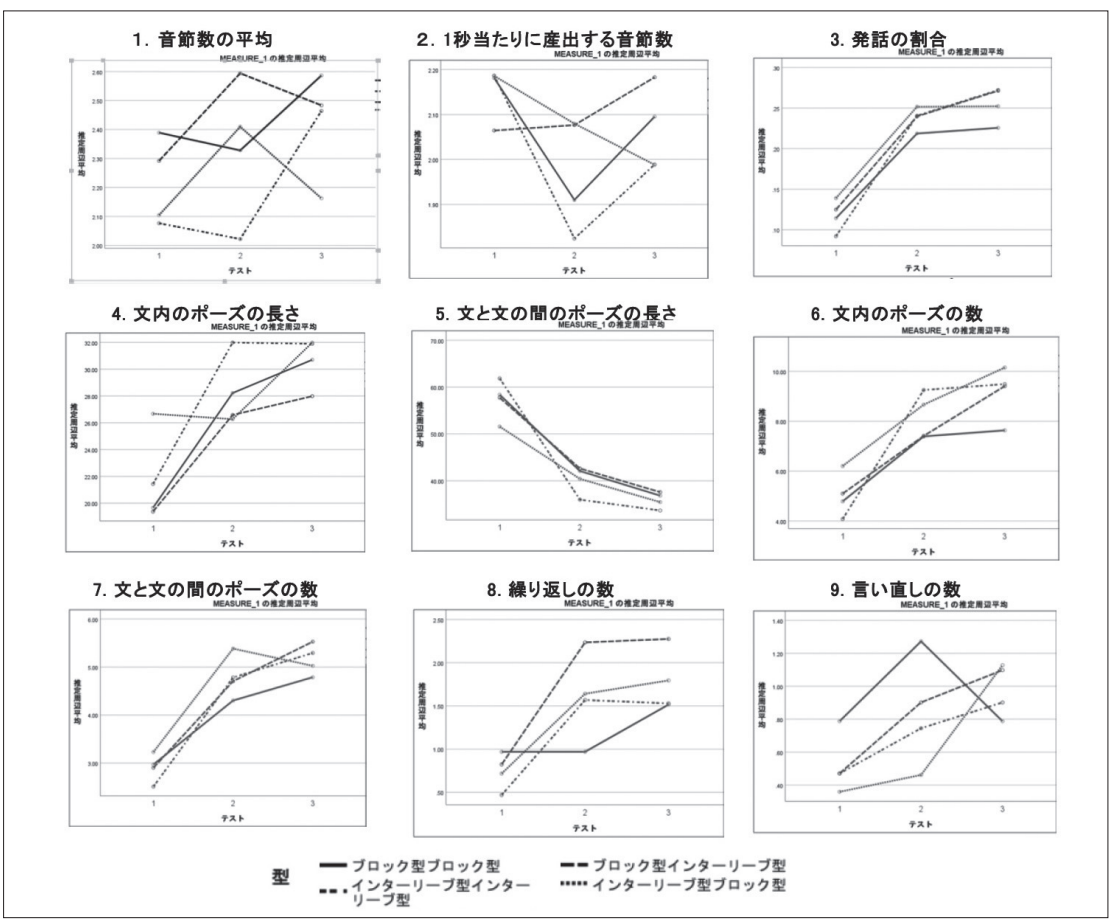
	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B 型	28	0.108	6.788	1	56	0.012	0.108(中)
I 型	30	0.025	1.421	1	56	0.238	0.025(小)

事前・中間・事後テスト

	n	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗
B B 型	11	0.055	1.555	2	53	0.221	0.055(小)
I I 型	17	0.053	1.471	2	53	0.239	0.053(小)
B I 型	17	0.111	3.305	2	53	0.044	0.111(中)
I B 型	13	0.093	2.708	2	53	0.076	0.093(中)

(注)偏イータ2乗内の( )の文字は効果量の大きさの程度の解釈を表す。

図2は，折れ線グラフを用いて9つの観点について4つの型(BB型，II型，BI型，IB型)を事前・中間・事後テスト間で表したものである。全体的な増加傾向が「発話の割合」(観点3)，「文内のポーズの長さ」(観点4)，「文内のポーズの数」(観点6)，「文と文の間のポーズの数」(観点7)，「繰り返しの数」(観点8)，となり，減少傾向が「文と文の間のポーズの長さ」(観点5)となった。また，「音節数の平均」(観点1)と「言い直しの数」(観点9)は型によって増加したり減少したりしていた。「1秒当たりに産出する音節数」(観点2)はBB型とII型は中間テストで減少し，事後テストで増加するV字型になっていたが，BI型は緩やかな増加傾向になり，IB型は中間，事後ともに減少していた。



■図2: 4つの型のグループ間のテストの比較

5 考察

RQ1は「日本人高校生がタスクの繰り返しを行った際に，初見の内容で，流暢性のどの側面が向上するのか」であった。そこで先ず，テストの効果量について本調査の結果を先行研究と比較していく。Suzuki（2021）は実験の結果，speed fluencyの「1秒あたりに産出する音節数」（観点2）とbreakdown fluencyの4観点全て（観点4~7）の合計5観点で，テスト間で効果量があったと述べた（speed fluencyの観点2が中程度，breakdown fluencyの観点4~7の全てで小程度）。一方，本調査では9つの観点の全て小程度以上の効果量があった（3, 5, 6, 7, 8が大，2, 4, 9が中，1が小）。これらのことから考えられる理由として，Suzuki（2021）の場合は，3回のタスクの繰り返しを協力者が各個人で行っただけだが，本調査ではタスクが1回終わるごとに教師による介入があり，そこで有効な表現方法等を指導したことが考えられる。毎回タスク後に書いてもらった「振り返りシート」では，多くの生徒が「I can see…. は使える表現だ」等，表現活動の有効性を書いていたことが印象的であった。

次にB型とI型についてだが，以下に述べる4つのことが考えられた。第一に，Suzuki（2021）では効果量のあった5観点で，いずれもB型の方がI型よりも良い結果であった。本調査でも「文内のポーズの長さ」（観点4）についてはSuzuki（2021）と同様，B型の方が短かくなっていたので，良い結果であったと考えられる。しかし，「文内のポーズの数」（観点6）については，Suzuki（2021）ではB型の方が回数が多かったのに対して，本調査ではI型の方が多かった。ただしいずれの型間の効果量も小程度なので，比較することは難しいと考える。第二に，「言い直しの数」（観点9）についてである。テスト間ではSuzuki（2021）では小程度の効果量も見られなかったが，本調査では小程度以上の効果量が見られた。また，B型の方が「言い直しの数」の効果量が中程度なのに対して，I型は小程度であった。数値を見てもB型の方がI型よりも発話中の「言い直しの数」が多かった。Saito et al.（2018）によると，repair fluencyは「発話時にどれだけ注意を払っているか」（p.596）ということになるので，B型の方が表現の確認をしながら発話をしていると考えられる一方，I型の方が3回タスクの繰り返しをするときに，毎回異なったマテリアルを使うので，発話時に応用がより利くと考えることもできる。その点で考えるのであれば，B型は同じ内容を繰り返して練習するので，決まった言い方は覚えやすいが，応用が利きにくく，初見の内容を描写するときに「言い直しの数」が増えてしまう傾向にあるのかもしれない。事後テスト後のアンケートでも「インターリーブ型はテキストが毎回かわるから応用をきかせないとむずかしかった。ただ文をおぼえるだけではだめだなと思いました」とコメントした者がおり，また似たようなこと書いた協力者が複数いた。第三に，全体的にSuzuki（2021）と比べて本調査の型間の違いの程度が小さかったことである。その理由として，協力者の学力や実験方法または学習意欲等が関係している可能性が考えられる。Suzuki（2021）ではCEFRレベルのA2からB1の大学生を対象にしたのに対し，本調査ではほとんどの協力者がCEFRレベルのA1であった。実験開始時には英語表現もあまり習得しておらず，何をどう言っていいのか分からないため，どの型でも絵描写タスク中に非常に長い沈黙をする協力者が多かった。また，使用マテリアルや録音時間も異なっていた。Suzuki（2021）では，6枚の絵を見て3分間で発話するのに対して，本調査では協力者の英語力を考慮した結果，Suzuki（2021）や森（2016）とは異なり，複数の絵ではなく1枚の絵を見て1分30秒で発話することとした。また，Suzuki（2021）では実験をするにあたり希望者を募ったが，本調査では実験に協力したい人を選んだわけではない。練習方法でもSuzuki（2021）とは異なり，本調査ではタスクを繰り返す際にペアを代えながら行ったため，マンネリ化は防げたのかもしれないが，学習そのものに興味を持っている者が多くなかったため，協力者はモチベーションが必ずしも高かったとはいえない。そして最後となるが第四に，テスト間での各型の効果量についてである。本調査では型間での違いについては効果量が小さいものが多かった半面，テスト間では効果量が大きいものが多かった（観点3, 5, 6, 7, 8の効果量が大である）ことを考えると，タスクの繰り返しをする場合は，英語習得レベルがそれほど高くない学習者を対象としたときには，基本的な語彙力や表現力が十分に身についていないことから考えても，B型やI型といった学

習方法で，どちらの型の方がより流暢性が向上するというよりは，タスクの繰り返しを何度も行いながら，その間で状況にあった語彙や表現等を指導することの方が学習者にとって効果的であると言えるのかもしれないことが示唆された。

次に、Suzuki（2021）の実験はタスクの繰り返しは3回であったが，本調査では更に3回追加して行ったので，その結果を考察する。結果からも分かるように，事前・中間・事後テストでは，それぞれの型の間で有意性は流暢性の9観点のいずれも認めることができなかったが，テスト間では9観点の全てで有意性を認めることができた。また効果量であれば，事前・中間テストの間では，型間では小程度の効果量が4観点（観点4, 6, 8, 9）だったところ，事前・中間・事後テストになると9観点中8観点（観点2~9）で小程度，そして1観点（観点1）で中程度の効果量がみられた。これは効果量はそれ程多くないものの，回数を増やすことによってCEFRレベルA1程度の学習者でも（cf. Suzuki（2021）ではA2～B1），B型とI型といった型での学習に効果が現れたということが言えるのかもしれない。また，テスト間では事前・中間テストで効果量が大きいものが5観点（観点3, 5, 6, 7, 8），中が3観点（観点2, 4, 9），小が1観点（観点1）となっている一方，事前・中間・事後テスト間では大が5観点（観点3, 5, 6, 7, 8），中が4観点（観点1, 2, 4, 9）となり，事前・中間テストと比べると，効果量の小程度がなくなり，中程度が増えている。これは若干であるが，型と同様に，長くタスクを行うことによってより一層流暢性の効果が現れてくる傾向になるかもしれないことを示唆している。具体的に効果量が大だったものを3つの流暢性の側面ごとにみると，先ずspeed fluencyである「発話の割合」（観点3：テスト時間全体の中で，ポーズなしに実際には話している時間の割合）の効果量が大きかった理由として，タスクの繰り返しをする中で，英語での表現方法に慣れたこと，またタスク後に表現方法の指導の介入が入り，有効な表現手段を得たこと，そしてタスクの繰り返しを6回することで，英語で表現することに自信をつけたこと等が考えられる。千菊（2018）でも3回のタスクの繰り返しを合計16回行い，事前・事後テストでその向上を測ったところ，speed fluencyである「1分間あたりの実発話語数」が伸びていた。また，本調査では，事後テスト後に取ったアンケートの中に，「最初はまったくしゃべれなかったけど，回を重ねるうちに沢山言えるようになったので成果が出て嬉しかったです！！」，「活動後にスピーキングの使える表現の練習をしたことで，かなり伝えることができやすくなったと思います。」や「たくさん話すことで英語力が上がったと感じた。」等，発話の向上が見られたコメントが多くあった。次にbreakdown fluencyについてであるが，「文と文の間のポーズの長さ」（観点5）がテストをするごとに長さが短くなる傾向にあり，「文内のポーズの数」（観点6）と「文と文の間のポーズの数」（観点7）がテストをするごとに数が増える傾向にあった。Saito et al.（2018）は，「文と文の間のポーズの数は，L2学習者が発話するときに何を言うべきか頭の中で整理できていることと関連している」と述べている（p. 596, 筆者訳）。そのことから，本調査ではタスクの繰り返しの練習をすることで，頭の中で発話する内容がより整理できるようになり，その結果「文と文の間のポーズの数」が増え，「文と文の間のポーズの長さが」が短くなったと解釈することができる。また，Saito et al.（2018）では「文内のポーズの数は発話者のその時の状態を表している」（p. 596, 筆者訳）とも述べられている。このことから，本調査でテストをするごとに「文内のポーズの数」が増えたのは，発話時により深く考えるようになったからではないかと推測される。ただし，考えながらするということは流暢に話すことにならないので，効果量が大きかったとしてもテスト間を比べ増加傾向であることは流暢性の点から考えると向上したとは考えにくいとも言える。最後にrepair fluencyの「繰り返しの数」についてだが，本調査ではテストをするごとに回数が増えていく傾向にあった。Saito et al.（2018）によると，repair fluencyは「発話時にどれだけ注意を払っているか」（p. 596, 筆者訳）になる。その解釈からすると，本調査で「繰り返しの数」が増える傾向にあったことは，タスクの繰り返しをすればする程，発話時により注意を払うようになると考えることができる。しかしこれも「文内のポーズの数」と同じで，増加傾向にあることは流暢性が向上したとは考えにくいとも言える。

RQ2は「タスクの繰り返しをする場合，どのような型の組み合わせが最も効果的に流暢性を向上させるのか」であった。Suzuki（2021）の研究では実験は3回のタスクの繰り返しを3回実施したが，本調査では更にタスクの繰り返しを3回増やし合計6回行った。また，B型とI型を1週間ごとに（3回ずつで）組み合わせ

ることで，どの組み合わせ（BB型・II型・BI型・IB型）が発話流暢性により効果的か調査した。ただし，効果量が大きくてもテスト間で増加傾向にある「文内のポーズの長さ」，「文内のポーズの数」，「繰り返しの数」についてはマイナスの効果があるかもしれないと考え対象外としたものと，そうではなく9観点の全てで効果量が大きかったものを効果があったと見做したものの両方の解釈で，それぞれの型が各観点の中でいくつ効果量が大きかったかを調べた（カッコ内は9観点全て含んだもの）。その結果，BB型が3（3），II型が5（7），BI型が3（5），IB型が2（3）となった。以上のことから，どちらの解釈からであっても，タスクの繰り返しを合計6回行い，事前・中間・事後テストで初見の内容を英語で描写させた場合，II型が最も発話流暢性の向上に効果的であると本調査の結果から分かった。これはB型の方が優れているというSuzuki（2021）の結果と異なるものとなった。ただし，型間の効果量を見る限り9つの観点中ほとんどが小程度なので（観点1が中程度で，それ以外の観点は小程度），前述した様に，学力のあまり高くない学習者ではB型やI型といったタスクの繰り返しではあまり差がつかないと言えるのかもしれない。また，練習後に介入を行ってしまったために，そのことでB型とI型のタスクの繰り返しの方法よりも介入の方が影響が大きくなってしまい，各型の差が大きなものにならなかったのかもしれない。その他の発見として，タスクの繰り返しを3回だけでなくさらに3回続けた場合，II型が最も効果的であったことから，全く同じことを3回繰り返すB型よりも，違う内容で3回行うI型の方が長期的な視野で見ると流暢性の向上に繋がるかもしれないということが本調査から示唆された。事後テスト後のアンケートでもB型とI型を比較して，「インターリーブ型のほうが話す内容を考えました」や「ブロック型のときより，インターリーブ型のときのほうが話せていたのではないかなと思います」，「より話す力がついたと感じたのはインターリーブ型です」等があった。

6 結論

本調査は，日本人高校生に対する短期間でより効果的なスピーキング活動である「タスクの繰り返し」について検証を行ったものだが，タスクの繰り返しを行った際に，初見の内容で，流暢性のどの側面が向上するのかについての回答として，「発話の割合」，「文と文の間のポーズの長さ」，「文と文の間のポーズの数」の3つの流暢性の観点の向上が期待できると言える。また，どのようなタスクの繰り返しの型の組み合わせが最も効果的に流暢性を向上させるのかについての回答として，タスクの繰り返しを3回×6回行うのであれば，複数の異なった種類のタスクを別の日に繰り返すI型を通して行うのが最も効果的であると言える。

限界点

1つ目として，本調査はB型とI型に分けてどちらがよりタスクの繰り返しを行った場合に大きな効果量が得られるのか調べた。しかし，本来であれば千菊（2018）のように実験群とは別に統制群も設けて調査を行うべきであった。しかし本研究は限られた時間に3クラスに依頼して行ったため，統制群を作って行うことは教育的配慮から難しかった。

2つ目として，事前・中間・事後テスト及び練習時に使用したマテリアルに対し，カウンターバランスを行わなかった点である。全てのマテリアルで英検準2級の問題を使ったので，回答レベルは同じであると思ったが，生徒の感想でも「今回は絵の情報量が少なかったので表現しづらかった」等があり，絵によってたくさん発話できるものとそうではないものがあることが分かった。今後タスクの繰り返しをする場合は，カウンターバランスも考慮して調査を進めたいと思う。

3つ目として，実験協力者の英語習得レベルを考えて，タスクの繰り返しの後に介入を行い，有効な英語表現や語彙の指導を行ったことである。そのために，その介入によって，どの程度調査に影響したかが分からなかった。しかし実際に教えれば教えるだけ生徒たちの表現力が付き，表情もよく活発にタスク活動に参加している姿を見るのは教師冥利というものであった。今回の調査結果でもタスクの繰り返しと

関連する指導の組み合わせに効果があったと考えられたため，さらにどのような組み合わせが最適かという問いで今後の研究を行うことが重要だと考えられる。

4つ目として，今回2週間タスクの繰り返しを行い，事前・中間・事後テストで効果を測定した。しかし，3週間以上行った場合にどうなるか，また事後テストの1か月後や半年後など，遅延テストを行った場合にどうなるか（指導の後の流暢性の維持）を調べることもできる。

5つ目として，今回は絵の描写のタスクのみを用いたことである。今後は意見陳述タスクや，生徒同士のやり取りを伴うディスカッションタスクなど別なタスクの繰り返しで同様の効果が見られるかを調べる必要がある。

教育的示唆

本調査では，タスクの繰り返しとその後の指導を行うことで流暢性が高まることが示唆された。その流暢性はB型とI型のどちらを行っても大きく異なるものではないことも示された。これに基づき，タスクの繰り返しとその後の指導を意図して授業内外で活動を取り入れていくことが必要だろう。その方法として，授業内であれば帯活動が効果的であると考えられる。1分間のタスクの繰り返しを3回するのであれば，ペア活動であっても15分以内に終えることは可能である。また，授業外であればSuzuki（2021）が行ったように，コンピュータに音声を吹き込んで練習することもできる。また，継続的に行うことで英語力も上がることが期待される。本調査では大きな違いは見られなかったが，Suzuki（2021）ではA2〜B1レベルの協力者で型間の差が出ていたことから考えれば，タスクの繰り返しを継続的に行い英語力を上げることで，より上級者だと型間に違いが生じ，より有効な型での指導が可能になるかもしれない。

教材に関しては，似たようなタスクを探す際は，英検の過去問題などを使うことが一つの方法だろう。本調査では英検準2級の問題を使ったが，生徒の学力に合わせて2級や準1級の問題を使うこともできる。級が上がるにつれてより発話内容も複雑になり，絵描写課題であれば2級以上では複数の絵を使ったストーリーのある描写が可能となる。

最後に，本調査で行った2週間（6回）というタスクの繰り返しの期間は，英検の一次の結果が来てから二次の面接の準備をするのに十分な期間である。本調査でも明らかとなったように，二次対策を行う上でもタスクの繰り返しは発話流暢性の向上に繋がり，短期間に流暢性を向上させるためには有効な方法であると考えられる。

謝辞

本研究を実施する機会を与えてくださり，また資料の使用を許可していただきました公益財団法人 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に厚く御礼申し上げます。とりわけ，指導助言者である小泉利恵先生には，本研究を実施，作成するにあたり，適切なご助言のみならず，いつも親身になっていただき，大変感謝しております。先生のおかげで，この研究を最後まで終えることができました。

本研究から得られたことを，これからの授業にも生かし，更に研鑽に努めるとともに，今後の英語教育の発展に貢献していきたいと思っています。

最後に，英検研究助成のことを紹介して下さり，また数々のご助言をしていただいた筑波大学の土方裕子先生，またゼミの皆さん，そして調査のために協力してくれた勤務校の同僚や生徒たちに深く感謝申し上げます。

引用文献

Ahmadian, M. J.（2011）. The effect of 'massed' task repetitions on complexity, accuracy and fluency: Does it transfer to a new task? *The Language Learning Journal, 39,* 269–280. https://doi.org/10.1080/09571736.2010.545239

Boersma, P. & Weenink, D.（2020）. PRAAT: Doing phonetics by computer [Computer program]. Version 6.1.14, retrieved 20 May 2020 from http://www.PRAAT.org/

千菊基司.（2018）. タスクの繰り返しを中心としたスピーキング指導と高校生の英語発話の質の向上.『日本教科教育学会誌』, 40(4), 25–37. https://doi.org/10.18993/jcrdajp.40.4_25

Date, M.（2017）. The Influence of Timing of Form Instruction During Practice Using Task Repetition on Task Performance. *ARELE: Annual Review of English Language Education in Japan, 28,* 97–112.

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning, 61*(2), 533–568. https://doi.org/10.1111/j.1467-9922.2010.00620.x

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354–375. https://doi.org/10.1093/applin/21.3.354

Fukuta, J.（2016）. Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research, 20*(3), 321–340.

Kahng, J.（2014）. Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. Language Learning, 64, 809—854. https://doi.org/10.1111/lang.12084

北原真冬, & 田嶋圭一.（2011）. 音声分析ソフトウェア PRAAT を用いた聴取実験: F_0 再合成による刺激作成と実験の制御（やさしい解説）.『日本音響学会誌』, 67(8), 3453–50. https://doi.org/10.20697/jasj.67.8_345

Koizumi, R., & In'nami, Y.（2014）. Modeling complexity, accuracy, and fluency of Japanese learners of English: A structural equation modeling approach. *JALT Journal, 36*(1), 25–46.

Koizumi, R., & Katagiri, K.（2009）. Changes in speaking performance of Japanese high school students: Longitudinal and cross-sectional studies at a *SELHi. ARELE: Annual Review of English Language Education in Japan, 20,* 51–60. https://doi.org/10.20581/arele.18.0_81

小泉利恵, & 栗嵜逸美.（2002）. 日本人中学生のモノローグにおけるスピーキングの特徴.『関東甲信越英語教育学会研究紀要』, 16, 17–28. https://doi.org/10.20806/katejo.16.0_17

Kormos, J.（2006）. *Speech production and second language acquisition.* NewYork, NY: Routledge.

松村昌紀.（2017）. タスク・ベースの英語指導—TBLTの理解と実践. 大修館書店.

文部科学省.（2017）. 平成29年度英語教育改善のための英語力調査　事業報告. https://www.mext.go.jp/a_menu/kokusai/gaikokugo/1403470.htm

文部科学省.（2018）.『高等学校学習指導要領解説 外国語編・英語編』開隆堂.

森敏行.（2016）. タスクの繰り返しが日本人高校生のスピーキングに与える注意焦点の変化, 太田信夫・佐久間康之（編）『英語教育学と認知心理学のクロスポイント（pp. 151–163）』北大路書房.

名倉秀人.（2010）.「PRAAT」を用いたニュース英語の音声分析. dialogos, 10, 101–114. https:/ /toyo.repo.nii.ac.jp/?action=repository_action_common_download&item_id=5099&item_no=1&attribute_id=22&file_no=1

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. Educational Research Review, 6, 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Saito, K., Ilkan, M., Magne, V., Tran, M., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low, mid and high-level second language fluency. *Applied Psycholinguistics, 39*(3), 593–617. https://doi.org/10.1017/S0142716417000571

Skehan, P.（1998）. *A cognitive approach to language learning.* Oxford University Press.

Skehan, P.（2003）. Task-based instruction. *Language Teaching, 36,* 1–14. https://doi.org/10.1017/S026144480200188X

Suzuki, Y.（2021）. Optimizing fluency training for speaking skills transfer: Comparing the effects of blocked and interleaved task repetition. *Language Learning, 71*(2), 285–325. https://doi.org/10.1111/lang.12433

Suzuki, S., Kormos, J., & Uchihara, T.（2021）. The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. The Modern Language Journal, 105(2), modl.12706. https://doi.org/10.1111/modl.12706

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Tavakoli, P., & Uchihara, T.（2020）. To what extent are multiword sequences associated with oral fluency?. *Language Learning, 70*(2), 506–547. https://doi.org/10.1111/lang.12384

Vercellotti, M. L.（2017）. The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied linguistics, 38*(1), 90–111. https://doi.org/10.1093/applin/amv002

資料1：タスクの繰り返しで使用したマテリアルの一例と一覧（英検準2級二次試験第3問から抜粋）

○タスクを終えた後の生徒の感想（「振り返りシート」より）

- 人がたくさんいることや行列ができるという表現は意外と使えそうなので覚えておきたいと思いました。
- 満員の店内の説明が難しい。
- 入れないという状況は説明できたけど,なぜこの状況になっているのかが言えなかった。
- 頭では日本語で考えているのに,英語で喋るからごちゃごちゃになる。
- まだ文章で伝えることができず,単語単語になってしまうので,次は調べたことをきちんと伝えられるようにしたいです。

マテリアル	使用資料
事前テスト	英検準2級二次試験第3問（2012年度第3回,問題カードA日程）
中間テスト	英検準2級二次試験第3問（2012年度第3回,問題カードB日程）
事後テスト	英検準2級二次試験第3問（2017年度第1回,問題カードB日程）
マテリアルA	英検準2級二次試験第3問（2018年度第2回,問題カードA日程）
マテリアルA’	英検準2級二次試験第3問（2018年度第2回,問題カードB日程）
マテリアルB	英検準2級二次試験第3問（2018年度第1回,問題カードA日程）
マテリアルB’	英検準2級二次試験第3問（2018年度第1回,問題カードB日程）
マテリアルC	英検準2級二次試験第3問（2017年度第3回,問題カードA日程）
マテリアルC’	英検準2級二次試験第3問（2012年度第1回,問題カードA日程）
マテリアルD	英検準2級二次試験第3問（2017年度第2回,問題カードA日程）
マテリアルD’	英検準2級二次試験第3問（2017年度第2回,問題カードB日程）
マテリアルE	英検準2級二次試験第3問（2017年度第1回,問題カードA日程）
マテリアルE’	英検準2級二次試験第3問（2012年度第2回,問題カードB日程）
マテリアルF	英検準2級二次試験第3問（2016年度第3回,問題カードA日程）
マテリアルF’	英検準2級二次試験第3問（2016年度第3回,問題カードB日程）

資料2：事前アンケート

英語に関するアンケート

3年　　組　　番　（氏名）　　男・女

1, 英語は好きですか?次の当てはまる箇所に〇をしてください。

好き　・　少し好き　・　あまり好きではない　・　嫌い

2, 次の中から英語で得意なものに〇をつけてください。（最低でも1つ,複数〇可）

読むこと　・　書くこと　・　聞くこと　・　話すこと

3, 英語のどんな力を特に付けたいですかですか?（複数〇可）

読むこと　・　書くこと　・　聞くこと　・　話すこと

4, 英検を取得している場合,合格した級に〇をつけてください。

2級　・　準2級　・　3級　・　4級　・5級

5, 海外へ行ったことがありますか。また行ったことがある場合その国と滞在期間
（例:アメリカに令和3年に1か月間）を教えてください。

ある　・　ない

行ったことがある国,期間:

6, いつから英語を習い始めましたか?（例:中学校1年生）

7, 英語について何でもいいので書いてください。（例:エッセー「英語と私」）

資料3：振り返りシート

10月　　日　　曜日

第　　回・振り返りシート

3年　　組　　番　（氏名）

1回目のペアワークを終えて

メモ

2回目のペアワークを終えて

メモ

3回目のペアワークを終えて

メモ

授業を終えて（今日学んだこと）

資料4：事後アンケート

事後アンケート

3年　　組　　番　（氏名）

1，今回のスピーキング活動は2週間にわたり，2種類の型（ブロック型とインターリーブ型）に分けて行いましたが，あなたは1週目と2週目でどの型で行いましたか。当てはまる方に○をつけてください。

1週目　　ブロック型・インターリーブ型
2週目　　ブロック型・インターリーブ型

2，以下の内容を参考にして，今回のスピーキング活動について思ったことや感じたことを書いてください。

- 1週目はテキストABC（A’B’C’）を，2週目はテキストDEF（D’E’F’）を使用
- ブロック型とインターリーブ型
- ペア活動
- ペア活動中の録音
- 振り返りシート
- ペア活動後にスピーキングの使える表現の練習
  （ I can see …, So I think he is in …, She wants to …）

注）ブロック型とインターリーブ型については，タスクの繰り返しが始まる前に黒板で丁寧に説明をした。
また，ブロック型とインターリーブ型は専門用語であるが，全体説明やタスク中に何度も使ったので自然と定着し，生徒の口からも出るようになった。

第35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

未知語推測トレーニングの効果:品詞への注目を促す明示的指導を通して

研究代表者:岡田　龍平　群馬県／関東学園大学附属高等学校　教諭
共同研究者:小室　竜也　茨城県／筑波大学大学院　在籍

《研究助言者:寺内　一》

概要

本研究では日本人英語学習者の未知語推測能力を高める2種類のトレーニングの効果を検証した。1つは品詞を同定する形式中心のもの，もう1つは未知語を推測するために必要な文脈的手がかりを同定する意味中心のものである。トレーニング前とトレーニング後の結果を比較したところ，時期による影響は見られたが，2種類のトレーニングの間には違いは見られなかった。この理由として，未知語推測には学習者の背景知識や熟達度など，さまざまな要因が関係していることが考えられる。

1 導入

1.1 研究の背景

英文を読解する際，生徒は分からない単語があると読むことを諦めてしまったり，逐一辞書を引いて長い時間をかけてしまったりすることが多い。一般に、読解中に未知語に遭遇した際には(a)読み飛ばす，(b)辞書を引く,(c)推測するという3つの方法を取ることができる。テキストのメインアイディアを理解するために不要である語であれば読み飛ばしても問題ないが，内容理解問題に直結する重要な語の場合には読み飛ばすことはできない。現在はスマートフォンを使って英文の写真を撮り，文字の読み取り機能を使ってDeepL(https://www.deepl.com/ja/translator)のような翻訳ツールにコピー＆ペーストすることも可能である。実際のビジネス現場でもメールやチャットの文面であれば翻訳ソフトウェアを使ったり，生成系AIで文法や英文スタイルの確認をしたりでき，児童・生徒が将来経験する英語の運用状況は今後さらに変化していくと考えられる。有用なツールが数多く利用できるようになったが，これらを使いこなすためには生徒自身の英語運用能力を高める必要がある。そのため，大学受験や期末テストの際にはAIツールや辞書の使用は許可されないことが多い。生徒は英文読解にあたって知らない単語に遭遇した際には，自分の語彙知識や背景知識を用いて推測しなければならない。生徒の未知語推測能力を高めることは，英文読解を自分の力でやり抜く力や、自律的な英語学習者の育成にもつながる。そこで、教師は生徒の未知語推測能力を診断し，品詞の同定や文脈的手がかりを同定するスキルを身につけさせる必要がある。そこで本研究では生徒の未知語推測能力の診断テストであるGuessing from Context Test（Sasao & Webb, 2018）を用いて，未知語推測に関わるスキルを伸長させるトレーニングの効果を検討する。

1.2 先行研究

多読のような英文読解やビデオの視聴などの言語活動を通して語彙知識を身につけていく方法は付随的語彙学習と呼ばれる。文脈を通した付随的学習は単語カードなどを用いた暗記や語呂合わせに代表される意図的学習に比べると、効率的ではない（Mondria, 2003）。しかし意図的学習と付随的学習は，車の両輪のように，英語学習においてはどちらも重要である。実際にNation（2013）の提唱する4 strandsでは，言語に焦点を当てた指導（意図的語彙学習）は、意味に焦点を当てたインプットや意味に焦点を当てたアウトプット，流暢性を伸ばす活動と同様に重要であるとされる。英文読解を通した語彙学習は意味に焦点を当てたインプットと流暢性の活動の両方で生じる可能性があるため，未知語推測能力が高いほど，この機会を効果的に利用することができると考えられる。未知語推測能力とインプットの間には循環関係が見られ，Perfetti and Hart（2002）のLexical Quality Hypothesisと共通する。Lexical Quality Hypothesisでは，語彙知識が豊富な読み手はテキストを深く理解することができ，インプット量が増えるほど語彙知識が豊富になるという好循環が提唱されている。未知語推測においても，推測に成功するとテキストを深く理解でき，テキストを深く理解できるとインプット量が増え，インプット量が増えると語彙知識が豊富となり，豊富な語彙知識は未知語推測に成功する可能性が高まる。ただし，未知語推測に成功する可能性は必ずしも高いわけではないため（25.6%: Nassaji, 2003; 12%–33%: Parry, 1991），未知語推測を通して語彙を学習するにあたっては学習メカニズムに影響を与える要因を適切に検討する必要がある。

文脈内で語彙を学習するメカニズムについて、Elgort et al.（2020）は母語研究のBolger et al.（2008）を引用してinstance-based memory modelを用いて次のように説明している：未知語の意味を推測しやすい文脈に複数回遭遇することにより，特定の文脈から語彙のコアである意味的特徴が抽出されるため，語彙の意味的表象が構築されると仮定される。語彙の意味的特徴とは，語彙意味論における意味成分のことを指す。例えば、manはHUMANでありMALEかつADULTである。一方womanはHUMANかつADULTであるがMALEではない。同様にboyとgirlはADULTではなくMALEかどうかで異なる。このように大文字で表される意味成分について，文脈を通した語彙学習では未知語のコアに対して累積的に蓄積され、最終的な意味表象につながる。

instance-based memory model以外にも付随的語彙学習のメカニズムは存在する。Jiang（2000）のrevised hierarchical modelでは，英語学習者のメンタルレキシコン（長期記憶に蓄積される辞書）として，英語の語彙が表す意味表象や概念はすでに母語で習得済みであることから，音韻や形態に関する知識を身につけていくと仮定する。例えば、boyという英単語を学習する際には，日本語で「男の子」という概念はすでに日本語のメンタルレキシコンに存在するため，英語のスペリングや発音を記憶し，形式と意味のマッピングを果たすことで，boyという語を習得すると考えられる。他にも、語彙連想課題（e.g., appleという語が与えられ，redやbananaを答える課題）の結果に基づき，語彙知識はコロケーションや定型表現をはじめとしたシンタグマティックな連想から上位語・下位語などのパラディグマティックな連想へと語彙知識は発達するという主張や（e.g., Eguchi et al., 2022; Fitzpatric, 2007），再認知識（提示された英語の意味を日本語の選択肢から選ぶ）から再生知識（提示された英語の意味を日本語に翻訳する）へと発達するという主張（González-fernández & Schmitt, 2020）もある。

生徒の文脈内語彙学習において，これらの語彙学習メカニズムの仮定が当てはまるのは未知語の意味を適切に推測した場合や，辞書を引いて適切な意味にアクセスできた場合であり，未知語を読み飛ばさないことが前提となる。生徒が未知語を読み飛ばさないようにする工夫として考えられるのは，テキストに占める既知語の割合を高め、推測しやすい文脈的手がかりを用意することである。テキストに占める未知語の割合が高いと十分な理解が得られにくく（e.g., Laufer & Rozovski-Roitblat, 2011; Van Zeeland & Schmitt, 2013），テキストに占める既知語の割合が95%から98%であることが求められている。生徒にはGraded Readerのように、自分の語彙サイズに合わせたインプットが必要となるだろう。

未知語の意味を適切に推測するためには文脈から得られる手がかりや十分な情報量が必要となる(Beck et al., 1983; Webb, 2008)。Beck et al.(1983)では英文を大きく4つに分類した。最も推測しやすい文脈はdirective context(When the cat pounced on the dog, he leapt up, and knocked down a shelf of books. The animals ran past Wendy, tripping her. She cried out and fell to the floor. As the noise and confusion mounted, Mother hollered upstairs, "What's all that commotion?")である。一定の意味範疇まで推測できる文脈はgeneral context(Joe and Stan arrived at the party at 7 o'clock. By 9:30 the evening seemed to drag for Stan. But Joe really seemed to be having a good time at the party. "I wish I could be as gregarious as he is," thought Stan.)と呼ばれる。これらの文脈は比較的推測しやすいが、全く意味を推測できないnondirective context(Dan heard the door open and wondered who had arrived. He couldn't make out the voices. Then he recognized the lumbering footsteps on the stairs and knew it was Aunt Grace.)や誤った意味へと誘導するmisdirective context(Sandra had won the dance contest and the audience's cheers brought her to the stage for an encore. "Every step she takes is so perfect and graceful," Ginny said grudgingly, as she watched Sandra dance.)は教育的には不適切な文脈であると言える。

未知語推測には既知語の割合や文脈情報量に加え，学習者自身の背景知識や言語的能力も必要となる。さらに，未知語推測そのものに関するスキルも重要となる。Nation(2013)は未知語推測スキルの構成概念に以下の3つを提唱している:(a)品詞を同定する，(b)文脈的手がかりを同定する，(c)実際に意味を推測する。これら3つのステップに関わる能力を診断するテストとしてGuessing from Context Test (GCT)がある(Sasao & Webb, 2018)。GCTは3つのセクションで構成され，1つのセクションにつき20問が含まれている。得られたそれぞれのセクションの点数は，4つのレベルに分けられ、学習者の未知語推測能力を診断することができる(表1)。

■表1: Sasao and Webb (2018) におけるGuessing from Context Testのスコアの解釈

Level	Raw score range			
	Rasch ability range	Part of speech	Discourse clue	Meaning
Advanced	Above 1 logits	16-20	16-20	16-20
Upper-intermediate	0 ～ 1 logits	13-15	11-15	11-15
Intermediate	-1 ～ 0 logits	10-12	6-10	6-10
Low	Below -1 logits	0-9	0-5	0-5

GCTのセクション1は品詞を同定する能力のテストである。品詞を同定する方法には「誰が何をしたのか」の指導が効果的である(Nation, 2013)。この指導では，主語や目的語には名詞が用いられ，動作は動詞で表されることを説明するものである。別のアプローチとして，派生形態素に着目する方法がある(Suzuki, 2018)。形態素には大きく，時制や人称による屈折形態素と品詞や意味を変える派生形態素がある。未知語の品詞を確認する手がかりとして重要となるのは-izeや-ty, -lyなどの派生形態素となる。GCTでは名詞・動詞・形容詞・副詞の4つを対象とし，その割合は9:6:3:2となっている。これらの語はNation(2006)のBNC word listにおける140,000語レベルの実在語を擬似語に置き換えており，接尾辞(屈折形態素と派生形態素の両方を含む)は実在語のものを使用している。接尾辞から品詞を同定できるのは20語中6語と設定されている。

GCTのセクション2は手がかりを同定する能力を測定する。GCTでは12種類の文脈的手がかりを採用し，50~60語のパッセージが作成されている。12種の分類は先行研究(Ames, 1966; Artley, 1943; Deighton, 1959; Dulin, 1970; Johnson & Pearson, 1984; McCullough, 1945; Seibert, 1945; Spache & Berg,

1955; Walters, 2006)に基づいている。具体的には以下の通りである：direct description（叙述的に直接説明されている：Their hutch was big enough for them to live in.), indirect description（明示的に説明されていないが間接的に説明されている：When it reaches the zenith at this age, it begins to go down.), contrast/comparison（反対の意味が示されている：Before she left she had gone into her kitchen, ostensibly to lock the back door, but in fact to look at her new food-mixer.), synonym（類義語が示されている：It is natural to feel better if you feel you look good; even in nature, cats preen themselves too.), appositive（次の文で言い換えられている：It is likely that the published works themselves are mutable: they might now look different.), modification（関係節で説明されている：The refectory where food was served was still a bit uncomfortable for her.), restatement（明示的ではないが次の文で言い換えられる：When she heard his words, her head lolled in the chair. She went to sleep immediately.), cause/effect（背景知識を使った因果関係を示している：The famous trees languished because of the terrible heat.), words in series（一連の語彙と関連する語彙で説明されている：The view was really beautiful as the light began to appear on the wide range of the sea; ahead, astern, and on either side of us.), reference（代名詞による照応関係で説明される：She had bought a new trilby for him. As soon as she gave it to him, he was very happy and put it on his head.), association（示される語句から連想を通して推測される：I averted my eyes from the sight of it.), example（例が示される：The show is held about various ways of having beverages such as tea and coffee.)。

未知語を推測する手がかりを同定する際の鍵となるのは前後のパラフレーズである。パラフレーズの種類として，未知語を上位語・下位語・類義語・反意語で置き換える方法がある（e.g., kettle → a container to boil water; sour → not sweet)。上位語や類義語などの語彙のネットワークは語彙連想課題ではパラディグマティックな連想と呼ばれ（Fitzpatrick & Thwaites, 2020; Fitzpatrick, 2007)，英文では語彙的結束性を高めるパラフレーズの方法として用いられる。一方、コロケーションのような統語的なつながりはシンタグマティックな連想と呼ばれる。未知語推測の手がかりとしては，パラディグマティックな語彙ネットワークの知識が重要である。語彙のネットワークに関する知識を用いて推測する方法はGCTではsynonym、cause/effect、word in series、association、exampleである。ネットワークの知識を活用するためには、順接や逆接など文脈の流れを適切に把握する必要がある。この流れを作る要因の1つにディスコースマーカーや接続詞が挙げられる。他にも、英文のトピックに関連した背景知識も手がかりを同定するために役に立つ。

未知語推測は形態素や語彙のネットワークの知識など，さまざまな要因が影響を与える複雑な活動であるが，学習者の熟達度がいくら高くなったとしても，未知語が全くなくなるとは考えにくい。自律的な英語学習者には未知語の意味を推測する能力が必要となることから，本研究では生徒の未知語推測能力を育成するトレーニングの効果について検討する。

1.3 本研究の目的

本研究では未知語推測のトレーニングの効果について検討することを目的とする。具体的な検証課題は以下の通りである。

形態素情報を用いて品詞を同定する形式重視の指導と語彙のネットワークを使って文脈の手がかりを同定する意味重視の指導は，どちらが未知語推測能力の伸長に効果的か？

2 方法

2.1 参加者

本研究に参加した協力者は，東京都内の私立大学に通う大学生41名であった。協力者は英語圏への長期の留学経験は有していなかった。アンケートによる英語資格試験に関する情報から，英語熟達度はCEFRでA1からA2レベルであると推定された。なお，トレーニングの介入条件は協力者間計画であるが，一般的な熟達度は等質であるように調整された。

2.2 マテリアル

未知語推測能力の測定にはGuessing from Context Test（Sasao & Webb, 2018）を用いた。GCTは3つのセクションで構成され，1つのセクションにつき20問，合計で60問の診断テストである。セクション1では品詞を同定する能力を，セクション2では未知語の意味を推測するために必要となる手がかりを同定する能力を，セクション3では実際に意味を推測する能力をそれぞれ測定する。

GCTではセクションごとに20点満点で集計し，学習者がどの程度の推測能力があるかを診断することを目的としている。そのため，本研究では3つのセクションを合計した60点を使用せず，セクションごとに点数が伸長したかを検討する。トレーニングの始まる1週間前と終了した1週間後の2回GCTを実施した。

2.3 トレーニング条件

本研究では形式を重視するものと意味を重視する2つのトレーニングを比較する。形式重視のトレーニングでは品詞の同定練習を行った。具体的には，形態素の知識に基づくもの（e.g., -izeは名詞のrealを動詞realizeにする，-lyは形容詞happyを副詞happilyにする）と辞書に記載されている品詞の記述を確認する練習を行った。一方，意味重視のトレーニングでは，ディスコースマーカーを使った語彙のネットワークの増強（e.g., howeverの後ろは反意語，for exampleは前に上位語で後ろに具体的な下位語）と背景知識を用いた連想知識の増強（e.g., 入院中の友達にお見舞いとして持っていくフルーツといえば…）を行った。

形式重視と意味重視のトレーニングは協力者間計画で1ヶ月行われた。トレーニングの回数は全部で4回であり，それぞれの内容は表2の通りである。

■表2: 未知語推測トレーニングの計画

	形式重視	意味重視
Week 1	品詞の種類と派生形態素に関する明示的指導（e.g., -izeは動詞，-lyは副詞）	ディスコースマーカーとパラディグマティックなネットワークに関する明示的指導（e.g., howeverは逆接、for exampleは例示，上位・下位語 [fruit-apple-orange]，反意語 [long-short]）
Week 2	形態素の知識に基づく品詞の同定練習	ディスコースマーカーに下線を引き，目標語の手がかりとなる語を同定する練習
Week 3	辞書に記載されている品詞の記述の説明	語彙のスキーマ・スクリプトに関する説明（e.g., キッチンで使われる語彙）
Week 4	辞書を引く前に品詞を確定させ，辞書を引く練習	背景知識を使った語彙のネットワークの増強（e.g., 入院中の友達にお見舞いとして持っていくフルーツといえば…）

調査協力者は授業における5分程度の時間でそれぞれのトレーニングに取り組んだ。Week 1およびWeek 3は教師主導の明示的な説明が行われ、Week 2とWeek 4はそれぞれ前の週の内容を踏まえた練習

問題が行われた。練習問題は30語前後の英文2題と学習者にとって未知語であると思われる語を使用した。形式重視のWeek 2では派生形態素を含む目標語が用いられ，意味重視のWeek 2ではディスコースマーカーを含む英文が用いられた。

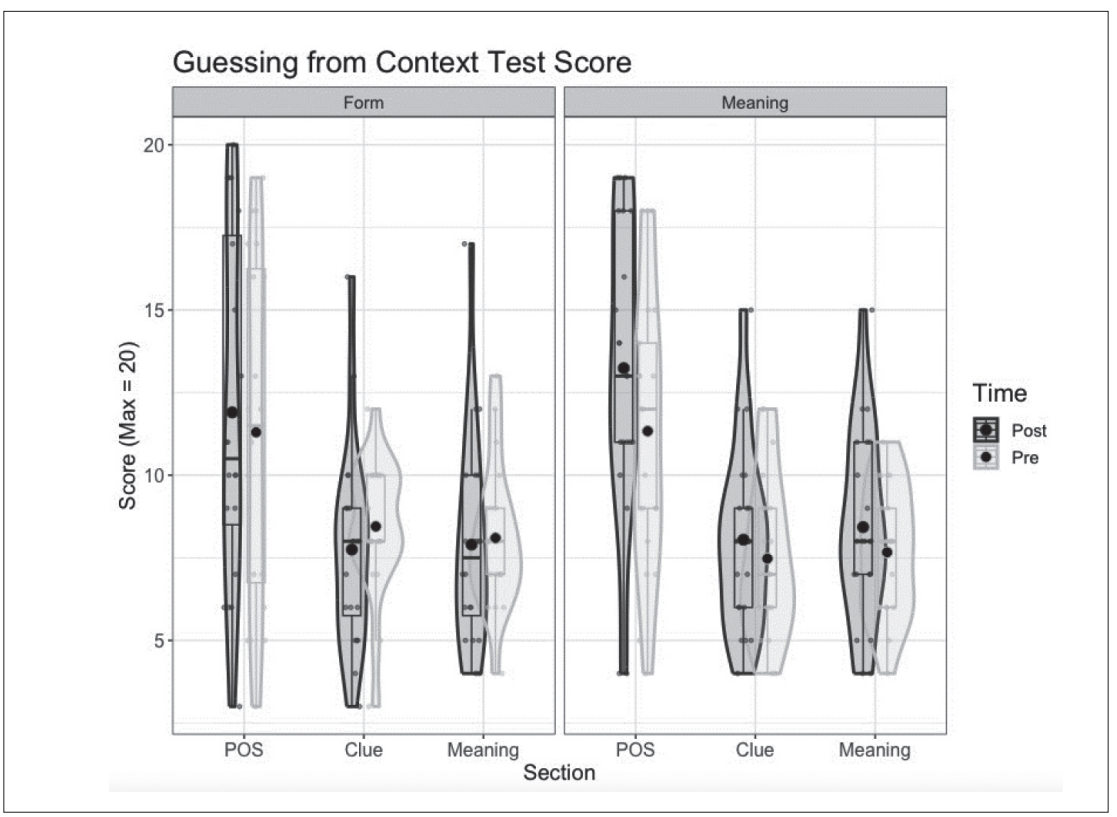
2.4 分析

Week 1の始まる前に実施したGCTのプレテストスコアと，Week 4の1週間後に実施したGCTのポストスコアを目的変数として使用した。テストの実施時期（pre vs. post）とトレーニング（形式重視 vs. 意味重視）を説明変数とし，ランダム効果として協力者を用いた線形混合効果モデリングを実施した。なお，GCTのスコアは3つのセクションごとに分けられていることから，モデルは全部で3つ作成された。分析にはR（ver. 4.2.3; R Core Team, 2023）とlme4パッケージ（ver. 1.1-34; Bates et al., 2015）を用いた。解釈のために説明変数はダミーコーディングが用いられ，交互作用は含まれていない。

3 結果

3.1 記述統計

記述統計は表3と図1に示す通りである。最も高かったのはセクション1の品詞同定能力であり，セクション2の手がかり同定能力とセクション3の意味推測能力の平均値に差はほとんどなかった。事前テストよりも事後テストの方が高い結果であった。



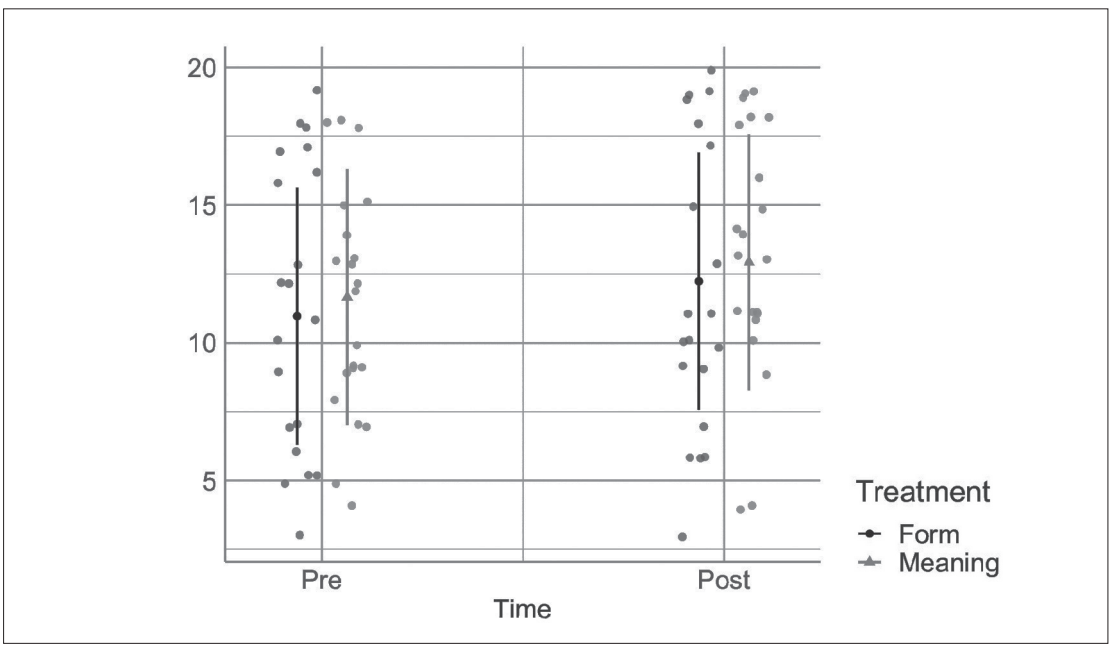
■図1：Guessing from Context Testのセクションおよび事前・事後テストごとのバイオリンプロット

■表3: Guessing from Context Testのセクションおよび事前・事後テストごとの平均と標準偏差

Time	Form		Meaning	
	Pre	Post	Pre	Post
Section	*M*(SD)	*M*(SD)	*M*(SD)	*M*(SD)
1	11.3 (5.24)	11.9(5.27)	11.33 (4.14)	13.24 (4.49)
2	8.45 (2.01)	7.75 (3.08)	7.48 (2.4)	8.05 (2.73)
3	8.1 (2.17)	7.9 (3.23)	7.67 (1.93)	8.43 (2.86)

3.2 セクション1:品詞同定能力

GCTのセクション1は品詞を同定する能力を測定している。セクション1のスコアを目的変数とした線型混合効果モデルの結果は表4および図2の通りである。



■図2: 品詞同定能力に対するテスト実施時期とトレーニングの影響

■表4: 品詞同定能力に対するテスト実施時期とトレーニングの影響

	固定効果				変量効果
パラメタ	推定値	95%CI	t値	p値	SD
切片	10.97	[8.88, 13.05]	10.47	< .001	18.69
テスト時期(事後)	1.27	[0.35, 2.19]	2.74	.008	
トレーニング(意味重視)	0.69	[-2.16, 3.53]	0.48	.632	

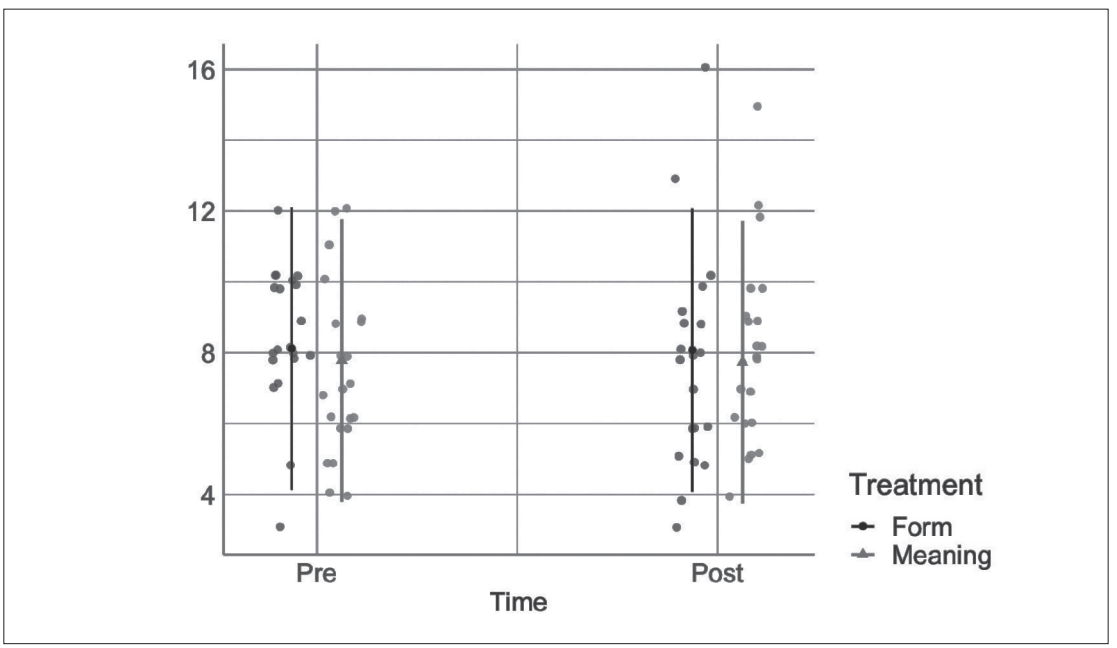
Note. formula: Score ~ Time + Treatment + (1 ¦ Participants)

推定の結果,事前テストに比べて事後テストのスコアは上昇したことが示されたが,トレーニングの効果の間には差は見られなかった。Marginal R2は.022であり,Conditional R2は.814であったことから,未知語推測トレーニングの効果は分散の2%を説明するにとどまり,参加者の個人差が非常に大きいことが

示された。これはトレーニングを実施中の英語授業によって品詞同定能力が伸長した可能性が考えられる。

3.3 セクション2:手がかり同定能力

GCTのセクション2は文脈からの手がかりを同定する能力を測定している。セクション2のスコアを目的変数とした線型混合効果モデルの結果は表5および図3の通りである。



■図3: 手がかり同定能力に対するテスト実施時期とトレーニングの影響

■表5: 手がかり同定能力に対するテスト実施時期とトレーニングの影響

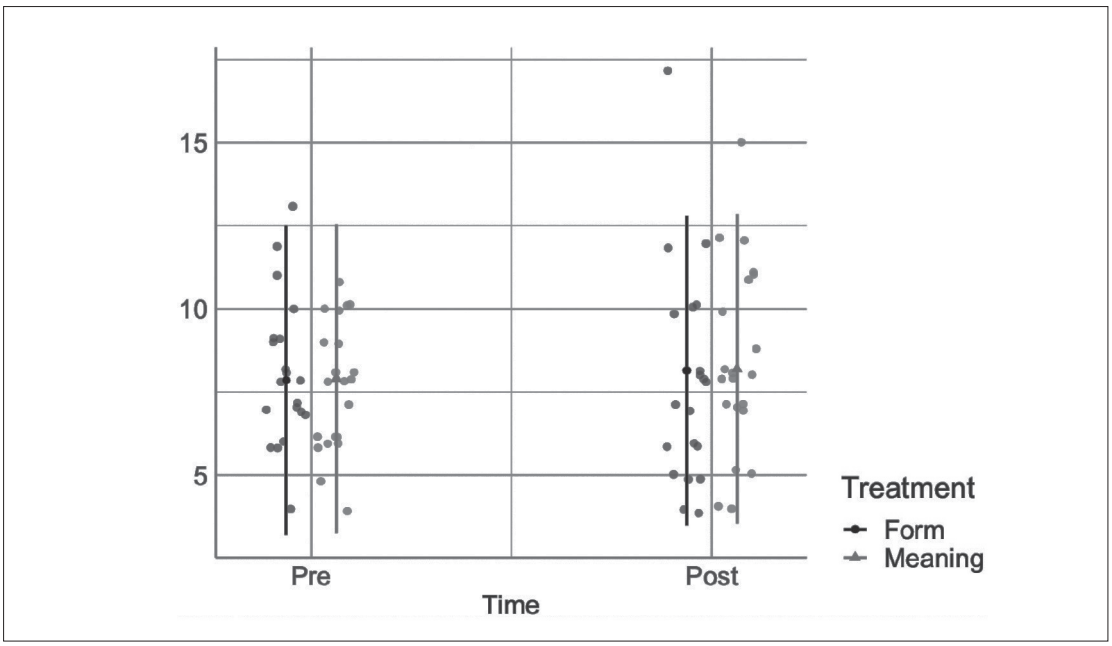
	固定効果				変量効果
パラメタ	推定値	95%CI	t値	p値	SD
切片	8.12	[7.05, 9.20]	15.11	< .001	3.72
テスト時期(事後)	-0.05	[-0.90, 0.80]	-0.11	.909	
トレーニング(意味重視)	-0.34	[-1.71, 1.04]	-0.49	.625	

Note. formula: Score ~ Time + Treatment + (1 ¦ Participants)

線型混合効果モデルの結果，テスト実施時期およびトレーニングの効果の間には差は見られなかった。つまり，トレーニングの前後で文脈から手がかりを同定する能力に変化があるとは言えず，形式重視と意味重視のトレーニングの効果についても結論を下すことができなかった(Marginal R^2 = .004, Conditional R^2 = .450)。

3.4 セクション3:意味推測能力

GCTのセクション3は実際に意味を推測する能力を測定している。線型混合効果モデルの結果は表6，図4の通りである。



■図4: 意味推測能力に対するテスト実施時期とトレーニングの影響

■表6: 意味推測能力に対するテスト実施時期とトレーニングの影響

	固定効果				変量効果
パラメタ	推定値	95%CI	t値	p値	SD
切片	7.85	[6.82, 8.89]	15.11	< .001	5.18
テスト時期(事後)	0.29	[-0.71, 1.29]	0.58	.562	
トレーニング(意味重視)	0.05	[-1.22, 1.31]	0.07	.940	

Note. formula: Score ~ Time + Treatment + (1 ¦ Participants)

混合効果モデルの結果，テスト実施時期およびトレーニングの効果の間には差は見られなかった。セクション2の手がかり同定と同様に，トレーニングの前後で意味を推測する能力に変化があるとは言えず，形式重視と意味重視のトレーニングの効果についても結論を下すことができなかった(Marginal R^2= .003, Conditional R^2 = .233)。

4 考察

転移適切処理(Morris et al., 1977)に基づくと、セクション1では品詞の同定に焦点を当てていた形式重視のトレーニングは，学習内容と測定内容が一致していることから，意味重視のトレーニングよりも効果的であると予測される。同様に，セクション2では文脈の中での手がかりを同定する意味重視のトレーニングが学習時と測定時の状況が一致しており，形式重視のトレーニングよりも効果的であると予測される。しかし，本研究ではセクション1および2でトレーニングの間に差が見られなかった。この理由として以下の2つが考えられる:(a)明示的な説明が多く実際の練習が少なかった，(b)GCT では擬似語が用いられている。

学習者の知識は大きく明示的知識と暗示的知識に分けられ，明示的知識はさらに自動化されたかどう

かの観点で別の構成概念をなしている(e.g., Suzuki, 2017; Suzuki & DeKeyser, 2017)。本研究では品詞を同定する明示的な説明とその練習を組み合わせていたため，自動化された明示的知識へと発達すると考えられた。一方，意味重視のトレーニングでは品詞に関する指導を行っておらず，事前テストから事後テストでスコアの変動(i.e., 自動化された明示的知識の発達)が生じないと考えられた。セクション1の分析の結果ではテスト実施時期の主効果が有意であり，意味重視のトレーニングも同様に品詞同定能力が伸長していた。なお，交互作用を含めたモデルでも時期の主効果が有意かつ交互作用は非有意であった。交互作用のモデルが含まれていないのは情報量基準が高かったためである。形式重視のトレーニングの効果が意味重視のトレーニングよりも優れていることが示されなかったのは，自動化された明示的知識の習得に至るために必要な練習量が少なかったことが原因であると考えられる。1回のトレーニングの時間が約5分であったことと，4回のトレーニングのうち明示的な説明が2回であったことから，練習の機会が少なかった。今後はより多くの練習の機会を設け，練習時間の影響を検討する必要がある。さらに，Data Driven Learningのような暗示的トレーニングの効果についても，明示的トレーニングと比較する必要がある。

2つめの理由として，形式重視のトレーニングでは派生形態素を用いた説明が含まれていたが，GCTでは擬似語(i.e., 実際には存在しない英単語)が用いられていた。つまり擬似語に派生形態素が含まれていない場合には，派生形態素を用いた推測方略を取ることはできなかった。具体的な擬似語の数は20語中14語である。同様に，意味的な連想関係でもシンタグマティックな連想関係については，学習者の熟達度が高かった場合には擬似語のコロケーションから推測が容易であった可能性がある。GCTでは受験者の既有知識の影響を最小限にするために擬似語が用いられていたが，本研究のトレーニングを適切に測定するためには不適切であった可能性がある。しかし，トレーニングの効果が全くなかったわけではない。セクション1の分析から，品詞の同定能力に関してはトレーニング前後で有意に能力が伸長していることが示されている。

セクション2とセクション3ではテスト実施時期もトレーニングも主効果が非有意であった。本研究で採用したトレーニングは5分の時間を4週間に渡るものであり，GCTで測定される未知語推測能力の発展に貢献しなかった可能性がある。今後は指導期間や練数回数の調整をはじめ，暗示的な指導や発見学習などの本研究で採用していなかった指導の検討し，形態素への気づきや読解法略の変化などのGCT以外の未知語推測に関する能力も多面的に検討する必要があるだろう。

またSasao and Webb(2018)のGCTは，学習者の未知語推測能力を連続値よりもむしろ4つのカテゴリカルな値として扱う方法も示されている。本研究では品詞同定，手がかり同定，意味推測の3つについていずれもintermediateの熟達度を有していることが示された。今後はupper-immediateやadvancedに属する学習者を対象とした研究も求められる。

謝辞

この研究の機会をくださった公益財団法人 日本英語検定協会，選考委員の先生方，特に貴重なアドバイスや励ましをくださった寺内一先生に心より感謝を申し上げます。また，研究に参加をしてくださった皆様に深くお礼申し上げます。

引用文献

Ames, W. S. (1966). The development of a classification scheme of contextual aids. *Reading Research Quarterly, 2*(1), 57–82. https://doi.org/10.2307/747039

Artley, A. S. (1943). Teaching word-meaning through context. *Elementary English Review, 20*(1), 68–74. https://www.jstor.org/stable/41382670

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

引用文献

Beck, I. L., McKeown, M. G., & McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal, 83,* 177–181. https://doi.org/10.1086/461307

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes, 45*(2), 122–159. https://doi.org/10.1080/01638530701792826

Deighton, L. C. (1959). *Vocabulary development in the classroom.* Columbia University Press.

Dulin, K. L. (1970). Using context clues in word recognition and comprehension. Reading Teacher, 23(5), 440-445. https://www.jstor.org/stable/20196343

Eguchi, M., Suzuki, S., & Suzuki, Y. (2022). Lexical competence underlying second language word association tasks: Examining the construct validity of response type and response time measures. *Studies in Second Language Acquisition, 44*(1), 112–142. https://doi.org/10.1017/S0272263121000164

Elgort, I., Beliaeva, N., & Boers, F. (2020). Contextual word learning in the first and second language: Difinition and inference error effects on declarative and nondeclarative knowledge. *Studies in Second Language Acquisition, 42*(1), 7–32. https://doi.org/10.1017/S0272263119000561

Fitzpatrick, T., & Thwaites, P. (2020). Word association research and the L2 lexicon. *Language Teaching, 53*(3), 237–274. https://doi.org/10.1017/S0261444820000105

Fitzpatrick, T. (2007). Word association patterns: Unpacking the assumptions. *International Journal of Applied Linguistics, 17*(3), 319–331. https://doi.org/10.1111/j.1473-4192.2007.00172.x

González-Fernández, B., & Schmitt, N. (2020). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics, 41*(4), 481–505. https://doi.org/10.1093/applin/amy057

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*(1), 47–77. https://doi.org/10.1093/applin/21.1.47

Johnson, D., & Pearson, P. D. (1984). *Teaching reading vocabulary.* Holt, Rinehart & Winston.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research, 15*(4), 391–411. https://doi.org/10.1177/1362168811412019

McCullough, C. M. (1945). The recognition of context clues in reading. *Elementary English Review, 22(*1), 1–5. https://www.jstor.org/stable/41382811

Mondria, J.-A. (2003). The effects of inferring, verifying, and memorizing on the retention of L2 word meanings: An experimental comparison of the "meaning-inferred method" and the "meaning-given method." *Studies in Second Language Acquisition, 25*(4), 473–499. https://doi.org/10.1017/S0272263103000202

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly, 37*(4), 645–670. https://doi.org/10.2307/3588216

Nation, I.S.P. (2013). *Learning vocabulary in another language.* Cambridge University Press.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of Functional Literacy, 11*, 67–86. http://digital.casalini.it/9789027296108

Parry, K. (1991). Building a vocabulary through academic reading. *TESOL Quarterly, 25*(4), 629–653. https://doi.org/10.2307/3587080

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sasao, Y., & Webb, S. (2018). The guessing from context test. ITL - International *Journal of Applied Linguistics, 169*(1), 115–141. https://doi.org/10.1075/itl.00009.sas

Seibert, L. C. (1945). A study on the practice of guessing word meanings from a context. *Modern Language Journal, 29*(4), 296–323. https://doi.org/10.1111/j.1540-4781.1945.tb00276.x

Spache, G., & Berg, P. (1955). *The art of efficient reading.* Macmillan.

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics, 38*(5), 1229–1261. https://doi.org/10.1017/S014271641700011X

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning, 67*(4), 747–790. https://doi.org/10.1111/lang.12241

Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension?. *Applied Linguistics, 34*(4), 457–479. https://doi.org/10.1093/applin/ams074

Walters, J. (2006). Methods of teaching inferring meaning from context. *RELC Journal, 37*(2), 176–190. https://doi.org/10.1177/0033688206067427

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*(2), 232–245. http://hdl.handle.net/10125/66826

B 実践部門 ● 報告Ⅳ ● 英語能力向上をめざす教育実践

中学校の「読むこと」における Mediation 活動を通した思考・判断・表現力の育成

研究者：林 さなえ　兵庫県／兵庫県立芦屋国際中等教育学校　教諭
（申請時：兵庫県／兵庫県立芦屋国際中等教育学校　教諭・兵庫県／神戸市外国語大学大学院　在籍）
《研究助言者：竹内　理》

概要

本実践は，Mediation活動により，読み物教材に対する中学生の思考力・判断力・表現力を高めることを目的とした。Mediationは，受容産出スキルとされ，他者と関わることで成立する。授業では，3年生20名がペアワークの形態で，別のペアとは異なる内容の物語文・説明文を読み，その内容を相手に伝える活動を行った。その中で，聞き手には話を聞きながら内容に関する質問を行うことを，また話し手には読んだ内容をもとにそれらの質問に答えることを求めた。

読み物教材の字義的な理解度は筆記再生テストで測定した。思考力・判断力・表現力については，生徒のやり取りを質的に分析した。その結果，Mediation活動は，読み物教材のジャンルに関わらず生徒の英文理解を促すことがわかった。さらに生徒のMediation活動のやり取りから，内容に関する質問や推論を活用した質問など，生徒が内容理解を深めるために辿った様々な過程が確認できた。

1 はじめに

文章理解とは，心の中に，テキストで描かれた内容に関する首尾一貫したイメージを作り上げることとされる（心的表象の構築；Kintsch, 1994）。そして，首尾一貫した心的表象を構築するためには，読解中に推論を行うことが必要とされる。推論とは，「メッセージの中で明示的に述べられていないテキストベースの項や命題を推論すること」（卯城，2009, p. 83）である。また池田（1994）は，読み手がテキストを理解するために，そこにはっきりと述べられていない内容を補って，既に得られている内容と統合させることと述べている。そのため教師は，読解中もしくは読解後に生徒が自らテキストの内容について推論をできるよう，「要旨をつかむ」「登場人物の気持ちを理解する」「文章中の最も重要な考えを判断する」といった読む目的を与えることが求められている。

しかし，首尾一貫した心的表象を構築できるよう，生徒に英語の読み方を教えることは，教師にとって最も難しい事項の一つでもある。多くの生徒は，読解中に初めて目にする語彙や構文を理解するのに苦労し，すべての文章を文字通り訳せないことに不安を覚える（Hamada & Takaki, 2021a, 2021b）。このような知識・技能面の問題から，生徒は読むことの本来の目的である「内容を理解する」ことに十分集中することができないでいる。さらに「授業では依然として，文法・語彙等の知識がどれだけ身についたかという点に重点が置かれ」（文部科学省，2017, p.6），思考力・判断力・表現力を高めるような言語活動が行われていないという課題がある。

そこで本実践では，Mediation活動によって生徒の思考力・判断力・表現力がどの程度向上するのかを

検証する。Mediation（媒介）は，CEFR補遺版（Council of Europe, 2020）で提示された従来の4技能とは異なる4つのコミュニケーションの形態（reception, production, interaction, mediation）の1つであり，reception, production, interactionを組み合わせたもので，“mediation make communication possible between persons who are unable, for whatever reason, to communicate with each other directly“と定義されている（Council of Europe, 2020, p. 34）。これは異なる言語間でのコミュニケーション（cross-linguistic communication）だけを想定したものではなく，知識や社会文化を伝達する際の媒介活動も含む。その方略として，例えば複雑な情報をわかりやすく伝えたり，相手の背景知識と伝える情報を結び付たりすることが求められる。そのようなMediation活動として，本実践ではリテリングのように，読んだ内容を話し手が聞き手に伝える授業を行った。話し手は，「読んだ内容を相手にわかりやすく伝える」ことを目的にすることで，聞き手が内容を理解しやすいように，何をどのように伝えるかに注意を払って伝えるべき内容を決定し，聞き手に伝えた。さらに，Mediation活動が通常のリテリングと異なる点は，聞き手は話し手の語る内容から，さらに知りたいことや疑問に思うことを質問する点にある。聞き手から質問を受けた話し手は，自分がテキストの内容を全て理解していないことを時には自覚しつつも，聞き手からの質問に答えなければならない。聞き手からの質問により，話し手は自分が伝えるべき内容をさらに掘り下げて考え伝える必要がある。そのため話し手のテキスト理解がより促され，内容を一貫して理解することにつながると想定される。このような協働により学習が促される仮説は，Vygotsky（1978）の発達の最近接領域理論（Zone of Proximal Development: ZPD）に基づいている。つまり，Mediation活動の重要な要素は，個人がそれぞれの役割を果たしながら，互いに協力することで共通の目的を達成する相互協力である。

このような理論的・実践的背景から，本実践の目的は，読解活動におけるMediation活動により英文読解における生徒の思考力・判断力・表現力を高めることである。生徒は，自身の語彙や文法知識の少なさを気にするあまり，読解における内容理解の優先順位が低くなっている場面が見受けられる。知らない単語や文型の意味を理解できれば満足するであろうが，読解活動の中で，自ら推論し，より深く文章を理解できるようになると，読解活動そのものが従来と異なるものとなる。Mediation活動は，読解活動のあり方を変えるきっかけになると期待される。

2 先行研究

2.1 読解における推論の役割

読む活動を行うことで，何かしらの痕跡（心的表象）が心の中に残る。そして，テキストの首尾一貫した心的表象を構築することが，読解が上手くできたことの証となる（Horiba, 2000）。理解度のレベルには，表層的記憶，命題的テキストベース，状況モデルという3つの段階がある（Kintsch, 1994）。表層的記憶とは，読み手がテキストにどのような単語が出てきていたかや，語順などの，言語的要素のみを記憶している段階を指す。命題的テキストベースの段階になると，「テキストそのものが何を伝えようとしているのか」（Grabe, 2009, p.46）が理解できているとされる。さらに，読者自身が自ら持っている背景知識と新たに得られたテキストの内容を統合することができると，最も深くテキストを理解できている状態（状況モデル）となる。状況モデルの段階に達するには，推論を行うことが不可欠である。これは，推論が，テキストに出てくる情報をつなぎ合わせ，またテキストと読者の背景知識とを結びつける働きを持っているからである（Horiba, 2000）。

推論は，橋渡し推論と精緻化推論に分けられる（Koda, 2005）。Clinton et al.（2020）による橋渡し推論と精緻化推論の違いは，推論によってどのように心的表象が形作られるのかということに関わる。橋渡し推論は，照応関係（Kenはheを表す）や，因果関係を表すものであるため，テキストを一貫して理解するために欠かせない。よって橋渡し推論が読解中に構築されなければ，内容理解に影響する（Singer et al.

1992, p.539)。一方の精緻化推論は，テキストの一貫した心的表象を構築するのに必ずしも必要とされない（Graesser et al., 1994）。しかし，精緻化推論には，文章のテーマについて考えたり，提起されている問題を解決することや，関連する背景知識を利用すること，著者のメッセージを批判したりすることが含まれる。従って，精緻化推論を利用して読解活動を行うことは必要なことと言える。

英文読解指導において，推論を利用して内容を理解させていくことが困難となる理由はいくつか考えられる。まず生徒が，文章理解に必要な単語の意味理解や構文解析を苦手としている場合がある。例えば，語彙知識が十分でない読者は，読解中に単語の意味を理解することに注意を払わなければならない（Grabe, 2009）。さらに，Grabe and Stoller（2020）は，「流暢な読み手は，単語が持つ基本的な文法情報を理解することで，節レベルで英文の意味を理解し，その内容をまとめて記憶することができる」（p. 18, 拙訳）と主張している。つまり，読解が困難な生徒にとって，読解中に起こる主な問題の原因は言語的な知識・技能によることが多い。語彙や文法に特段の注意を払いがちな生徒にとって，適切な指導を行わなければ，テキストを深く理解する機会が失われることとなる。

推論生成に影響を与える2つ目の要因はテキストジャンルである。例えば，物語文は説明文よりも理解しやすいとされる（Horiba, 2000）。物語文は一般的に，登場人物や場面設定といった，いつ，どこで何が起こるのか，また登場人物が起こす行動とその理由や結果が描写されるなどといった，共通の構造を持っている。ゆえに，物語文を普段から読んでいる生徒は，初めて読む物語文に出会っても，次にどのようなことが起こるかを予想することができる。一方，説明文を読む目的は新しい知識を得ることである（Ushiro et al., 2015）。Grabe（2009）は「説明文読解では，背景知識や個人の経験と，テキスト内容との関連は少ないことがよくある」（p. 250, 拙訳）と述べている。つまり，説明文を読む多くの生徒には，読んでいる話題に関する知識が十分でない場合，テキストを読みながら，新しい知識を学習することが求められる。これは生徒にとって大きな負担となる。このように，推論によってテキスト内容を深く理解できるかどうかは，生徒がテキストに関する十分な背景知識を持っているかによる影響が大きい。

生徒が英文読解中に自分で推論できるようになるには，明確な読む目的を与えることが重要とされる。van den Broek et al.（2001, p. 1082）は，読み手は，読解前に与えられた目標に向かって，文章を理解しようとすることを述べている。また，Ushiro et al.（2018）は，生徒は読む目的により読み方を変えていることを明らかにした。そこで本実践では，読む目的を与えるため，タスクベースのインストラクションを用い，生徒の読み方を誘導することとした。

2.2 読解後の言語活動としてのMediation

2.2.1 目的を持った英文読解におけるMediationの役割

Mediationは，CEFR補遺版（Council of Europe, 2020, p. 35）において，背景知識と結び付けること，わかりやすい言葉で言いかえること，複雑に絡み合った情報を紐解き分かりやすく伝えること，例を出し詳しく述べること，キーワードを強調するなどして明確に伝えることなどが挙げられている。前述のMediationを含めた4つのコミュニケーション形態におけるreceptionは，従来の4技能では「聞く」と「読む」の技能に，そしてproductionは，「話す」と「書く」の技能に分類される。North and Piccardo（2016）は，interactionは，receptionとproductionの集合体であり，情報の受信者と発信者が意味を共同構築することと述べている。そこでは情報の交換が行われ，その情報を明確にするための内容も含まれる（Council of Europe, 2020）。一方のMediationは，interactionと似た特徴を持ちあわせているが，reception, production, そしてinteractionとも区別されている（Coste & Cavalli, 2015）。Interactionがコミュニケーションの参加者同士の関わりから発生する相互作用であることに対し，Mediationは個人的または社会的なつながりを構築するものとされている。Coste and Cavalli（2015）は，「Mediationは，その場に参加している者がお互いの距離を縮めるため，段階を踏んで，様々な手段を講じ，行動に移すこと」（p.27, 拙訳）と述べている。

さらにMediationでは，interactionのように情報やメッセージが伝達される一方で，Mediator（情報の媒

介者）の解釈も併せて受け手に伝達される。Mediatorの役割は，受けた情報をそのまま発信することでなく，受けた情報とその情報の受け手とを結びつけることにある。よって，Mediatorから伝えられる新たな情報は原文と同じではなく，情報の受け手の理解度は，Mediatorによって左右される。Stathopoulou（2015）は，Mediatorには，伝える目的やその環境に応じて，原文からどのような内容を伝えるべきかを決定する必要があることを強調している。このように，Mediatorによる，情報を受け，その内容を吟味し，受け手の理解に合わせて自分の言葉で伝えることは，interactionで行われる情報交換とは大きく異なっている。

North and Piccardo（2016）は，Mediation活動を展開していくにあたり，説明すること，要約すること，異なる言語で書かれた文章の内容を明確にすること，そして理解を促すこと，などを用いたいくつかのアイデアを紹介している。なお上記の行為は全て，私たちが日常生活の中でごく自然に行っていることである（Council of Europe, 2020）。例えば授業中に，生徒がテキストの内容を理解しやすいように，教師がMediatorとなって生徒の理解度に合わせて扱う内容を様々な方法で取り入れることや，生徒がテキストの内容を理解できない場合，教師がテキストには書かれていない補足情報や説明を加え，生徒が授業の内容を理解しやすいよう工夫を加えることも身近なMediationの例と言えよう。教師がMediatorとしての役割を果たすことで，生徒はテキストの内容と自分自身とのつながりを持つことができ，よってより深くテキストの内容を理解できることにつながる。

2.2.2 Mediation活動とリテリング活動の違い

Mediationはテキストとの媒介行為や共同的なやり取りを通した考えの形成なども含み（投野・根岸，2020），複数の異なる技能が同時に必要とされる。Mediation活動において話し手は，読んだテキストの内容や理解したことのメッセージを伝え（Coste & Cavalli, 2015），さらに話し手自身の解釈も同時に聞き手に伝えている。一方で聞き手は，聞いた内容に対して質問をすることで，話し手の内容理解を深める手助けをする。したがって，Mediation活動は，話し手と聞き手が互いのテキストへの内容理解を促す必要がある。

目的を持った読解活動は，推論生成や一貫した読みといった読者のテキスト理解を促す（Horiba, 2000; Ushiro et al, 2018）。例えば，授業でよく行われているリテリングの目標は，「テキストを読んで理解したことを，一貫して相手に説明すること」であるため，リテリングはテキストを理解させるための効果的な指導法とされている（Gibson et al., 2003）。

リテリングは通常2人1組で行われ，話し手と聞き手に分かれて，生徒が主体となって行われるポストリーディング活動である（卯城，2009）。また聞き手は話し手にフィードバックを与える必要はない（Koskinen et al., 1988）。Kai（2009）は，リテリングが生徒の文章の理解度を高めることを実証した。また，リテリングを行った後の生徒は，テキスト中の重要な情報とそうでない情報を区別できるようになっていたことも明らかになった（Kai, 2008）。

しかし，リテリングには注意しなければならない点がある。第一に，L2のアウトプット能力が十分にある生徒にとってリテリングは有益な活動であるが，そうでない生徒にとっては，語彙や文法などの処理が追い付かず，リテリングが必ずしも有効になるとは限らない点である（Kai, 2008）。仮にそのような生徒がリテリングに取り組んでも，テキストの意味を聞き手に説明するのではなく，テキストの単語を記憶するだけになってしまう恐れがある。第二に，Koskinen（1988）は，聞き手となる生徒に役割を与える必要があることを示唆している。具体的な方法として，ワークシートを用いて話し手の理解度を確認したり，話し手に感想を述べたりすることが挙げられる（卯城，2009）。これらの方法により，聞き手が単に「話を聞くだけ」という状況は避けられる。一方，本実践では，これらの問題を解決するため，話し手と聞き手による協働的な方法で読解を促すことが期待されるMediation活動を試みる。

使用する言語は，母語もしくはL2で行う方法が考えられる。卯城（2009）は，外国語でリテリングを行うことは，生徒への負担が大きく，かつ対象とする言語のアウトプット能力と関わるため，読解力を問う場合には母語を用いることを提案している。また，甲斐（2008）は，リテリングの活動は母語（日本語）で実施

する方が，外国語（英語）で実施するよりも，情報の重要度に関わらずより多くの内容を伝えることができることを明らかにした。本研究の使用言語はMediation活動を含め，原則は日本語で行うこととした。しかし担当クラスには日本語が母語でない生徒も含まれており，また内容によっては英語の方がスムーズに出てくる生徒もいることから，L2（英語）も使用可とした。

2.2.3 Mediation活動の形式

Mediation活動にはMediating a text（取引・交渉言語使用），Mediating concepts（評価・問題解決言語使用），そしてMediating communication（創造・対人言語使用）という3つの形式がある。Mediating a text では，言語的，文化的，意味的そして技術的なことが理由で，そのテキストにアクセスできない相手に，その内容を伝えることを軸としている。これには，特定の情報を伝えること，データの説明を行うこと，文章を翻訳すること，メモを取ること，意見を言うことなどが含まれている。Mediating communicationでは，個人的，社会文化的，社会言語的に異なる背景を持つ参加者全員が，お互いによい影響を与える関係を築きながら，円滑なコミュニケーションを行えることを目的としている。Mediating conceptsにおいては，他者が知識や概念にアクセスしやすい過程について述べている。そのためNorth and Piccardo（2016）は，Mediationを通じて知識や概念とのつながりを深める方法として，グループでの協働と，グループワークを先導していくことの2つを述べている。これらの方法では，参加者がinteractionを通してテキストの意味を解釈したり抽出したりするために，互いに協力する必要がある。

Council of Europe（2020）では，Mediation活動のためのCan-Doリストが作成されており，熟達度によって，Can-Doリストにある能力記述子の内容は異なる。例えば，協働で意味を解釈する場合，A2レベルの聞き手は，適切な質問を投げかけることで，話し手が内容をよく理解できるようにすることが期待される。一方，B1レベルの聞き手には，話し手が言ったことを繰り返し，そして確認することが期待されている。前述したようにMediationは協働作業であるため，Can-Doとして達成すべき内容が違っていても，お互いに助け合いながら，共に内容理解を深めていくという考えに変わりはない。

本研究では，Mediation活動の内容を，生徒の文章理解度向上と密接な関係があると想定されるCollaborating to construct meaningとした。そして，Can-Do的な目標として，B1レベルの生徒は，「相手の言ったことを繰り返して相互理解を確認し，内容理解を促すことができる」，A2レベルの生徒は，「適切な質問をして，相手が自分の言いたいことを理解しているか確認できる」を選んだ。これは本実践を行う生徒の英語力が，個々人で大きく異なっていたためである。読解に対する不安を軽減し，幅広い生徒のレベルにあわせて目標を設定した。

2.2.4 研究課題

Mediationの教育効果はVygotsky（1978）の理論体系に位置づけることができる（North & Piccardo, 2016）。VygotskyのZPDには，子どもは他者（親，教師，友人など）の助けを借りて能力を高め，問題を解決することができるとある。このように，ZPDは，協働活動を通して，支援を受けることで得られる発達の領域を指す（Thorne, 2000）。本研究のMediation活動では，聞き手の質問をZPDにおける支援と位置づけることで，協働作業を促進させ，より深いテキスト理解につながることが期待される。しかし，Ohta（2000）は，支援が多すぎたり，提供される課題が簡単すぎたりすると，生徒が主体的に取り組む機会が奪われるため，ZPD領域の拡大が見込めないことを指摘している。そこで本実践では，以下の研究課題の検証を通して，生徒が自ら推論を行うための活動としてMediation活動の可能性を探る。

RQ1	Mediation活動により，英文の字義的な理解度はどの程度向上するか。
RQ2	Mediation活動において，生徒はどのような思考力・判断力・表現力を働かせているか。

3 方法

3.1 参加者

参加者は，代表者が授業を担当している中学3年生の20名である。生徒たちには研究の内容を説明したあと，この研究を授業内で行うこと，映像やワークシートも含め，教員以外が使用しないこと，また活動中の個人名も特定されることがないように分析することを口頭で説明し，自由意志での参加の同意を得た。また中学3年進級時で英検2級合格者が7名，準2級合格者が2名，3級合格者が3名（それ以外の生徒は英検未受験）と，生徒たちの英語運用能力も平均より高い傾向にあった。

3.2 教材

本実践では，英検2級過去問の読解問題を用い，物語文と説明文を4つずつ使用した。テキスト選定においては，言語的な難易度と題材という2つの基準を考慮した。特に題材については，背景知識の個人差による文章理解度への影響を避けるため，生徒が使用している教科書には掲載されていない題材を使用した。例えば物語文では，歴史上の人物について書かれているもの，説明文では人間科学や動物科学に関するものを題材とした。

具体的に，物語文として，ダーウィンのように進化論に着想を得た人物を描いた“Alfred Russel Wallace”（以下，“Wallace”），第二次世界大戦中にアメリカ政府の資金援助を受けて巨大な飛行機を作り，飛行させた人物を描いた“The Sky Is the Limit”（以下，“Hughes”），「タイムスタディ」を導入し近代工場を発展させた人物を描いた“Fredric Winslow Taylor”（以下，“Taylor”），そして大陸移動説を発見したが脚光を浴びる前に亡くなった人物“Floating Continents”（以下，“Wegner”）の4つを選んだ。説明文としては，カナダの厳しい冬を生き抜くカエルとその秘密と，そのカエルの特異な性質を利用し人間の臓器移植に応用する方法を紹介した“An Incredible Survivor”（以下，“Frog”），塩分の摂りすぎが深刻な健康被害をもたらす危険性を説いた“Too Much Salt”（以下，“Salt”），左利きが少ない理由と左利きの利点を説明した“Left-Handed Competition”（以下，“Left-Handed”），そして人間だけでなく動物も白内障になる経緯とその治療法を説明した“The Gift of Sight”（以下，“Eyesight”）の4つを選んだ。

選定したテキストが物語文のジャンルに属するか否かを確認するために，物語文法の考え方を適用した。Thorndyke（1977）によると，「すべての物語に必要な構成要素は，設定，テーマ，プロット，問題解決である。」（p. 80）。設定とは，時間，場所，主人公に関する情報，テーマとは，主人公が達成したい話を表し，またプロットとは，主人公の望む目標に関連するエピソードが書かれており，問題解決とは，テーマに基づいたストーリーとその結果である。例えば“Wallace”では，時間，場所，主人公の情報など，設定に関する要素がすべて冒頭で与えられていた。テーマについては，この物語の主人公であるWallaceが，進化論を証明するという目的を達成するために奮闘する姿が描かれていた。その具体例として，Wallaceが進化論で有名なDarwinに手紙を送るなど，彼の行動が物語に盛り込まれており，これをプロットとして判断した。物語の最後は，WallaceではなくDarwinのみが「進化論」で有名になったことが述べられているが，Wallace自身が有名にならずとも，彼の目標は達成されたことがわかる。このようにして，4つのテキストに物語文を構成する要素が全て含まれていることを確認した。

説明文の特徴は，物語文のものとは異なる。Ray and Meyer（2011）は説明文の構造的な複雑さを指摘した。そして，説明文の最も顕著な特徴は，その階層的な構造であると述べている。説明文では，まず上位の新しい概念が紹介され，次に詳細が説明される。読み手が読んだ内容を書き出すとき，詳細な情報は文章の主題の後に想起される（Schallert & Tierney, 1980）。つまり，読み手は説明文の内容を理解するために，

まず上位の情報を理解しようとする傾向がある。例えば，本実践で扱った説明文“Frog”では，まずカナダに驚くべき生態のカエルが存在することを紹介し，そのカエルがカナダの厳しい冬をどのように生き延びているのかを，詳細に述べている。その後の本文では，カエルの特性を生かした臓器移植について言及している。このような階層構造が4つの説明文に含まれていることを確認した。

テキストの難易度は，CEFR-J B2.1～B2.2レベルに設定されているかを確認するため，CEFR-Based Vocabulary Level Analyzer（Uchida & Negishi, 2018）で計算した（表1参照）。

■表1: 使用教材のレベル

テキスト	言語的指標				CEFR レベル
	FKGL	VperSent	AvrDiff	BperA	
Wallace	9.39	3.14	1.70	29%	B2.1
Hughes	9.07	3.18	1.76	29%	B2.2
Taylor	10.19	3.38	1.78	29%	B2.2
Wegner	10.08	3.25	1.82	32%	B2.2
Frog	8.73	3.43	1.66	24%	B2.1
Salt	9.88	3.42	1.83	27%	B2.2
Left-Handed	10.63	2.81	1.87	33%	B2.2
Eyesight	9.84	3.24	1.81	30%	B2.2

注. FKGL = Flesch-Kincaid Grade Level, VperSent = 一文あたりの動詞の数, AvrDiff = A1を1, A2を2, B1を3, B2を4（C1とC2レベルの語彙は計算に含めない）とした場合の語彙の困難度の平均, BperA = Aレベルの語彙に対するBレベルの語彙の割合

3.3 手順

表2に示す手順で実践を行った。1･2回目は長文読解後，Mediation活動を行わず筆記再生，推論質問に取り組んだ。一方で3･4回目は長文読解後にMediation活動を行い，引き続き筆記再生，推論質問に答えた。1･3回目は物語文，2･4回目は説明文を使用した。

■表2: Mediation活動の手順

Mediationなし（1･2回目）	Mediationあり（3･4回目）
1. 物語文･説明文を読む（8分）	1. 物語文･説明文を読む（8分）
2. 筆記再生テスト（8分）	2. Mediation活動（1チーム5分×2チーム 計10分）
3. 推論発問テスト（12分）	3. 筆記再生テスト（8分）
	4. 推論発問テスト（12分）

Mediation活動は，4技能の全てを含めた活動であるため，より日常のコミュニケーションに近いものとしてCEFRでも位置づけられている。Mediation活動を行わない場合と比べて手順が多くなるが，Mediation活動はコミュニケーションを取っていく中で，話の内容を整理し，理解度を高めることを目的としているため，どの程度話が理解できるようになったかを測るものとして適当であると考える。

読解活動では，生徒たちにはいずれの時も，物語文･説明文を黙読するよう指示した（8分）。その後，日

本語または英語のどちらか好きな言語で記憶したことを書き出すよう伝えた(12分)。それから推論問題に答えるよう指示した(12分)。Mediation活動では, 4名のグループをA／Bの2チームに分け, 各チームのMediation活動に5分間を与えた。詳しい手順として, まずチームAが, 5分間で自分たちが読んだ内容について, チームBの生徒たちに話した。チームBは, チームAの話を聞きながら, その内容をより深く理解するために, チームAの話の中でさらに知りたいことについて質問した。その後, 役割を交代し, チームBが読んだ内容を語り, チームAはチームBが語った内容に対して質問をした。なお, チームAとチームBにはそれぞれ異なるテキストが割り当てられるようにした。

Mediationの様子は, iPad(Apple社製)を使って録画し, その後の分析のためにアンケート調査を実施した。タブレットは各グループの近くに配置し, 生徒たちの音声や体の動きを捉えた。その後, 録画をもとにMediation活動での生徒のやり取りを書き起こした。生徒に配付したアンケートでは, ①Mediation活動を楽しめたか, ②Mediation活動は役に立つ活動であったか, という2つの質問をし, 5は賛成, 1は反対の5段階で評価させた。さらにMediation活動に対して「よかったこと」,「難しかったこと」,「感想」を自由に記述させた。

3.4 採点とデータ分析

3.4.1 筆記再生の統計的分析

筆記再生テストの採点として, まずは, 物語文と説明文のテキストをCarrell(1985)のアイデアユニット(IU)に分割した。1つのアイデアユニットには節が1つあり, 不定詞および動名詞節や接続節は別のユニットとみなした。2名の評価者がこれらの分割を行い, 両者の平均一致率は物語文で92.00%, 説明文で92.25%であった。その後, すべての生徒の筆記再生プロトコルを, 2名の評価者が採点した。採点では, 各IUに対応する意味が再生されていれば, 1点を与えた。採点の一致率は, 物語文で89.25%, 説明文で86.75%であった。不一致点は話し合いで解決した。

統計的な分析として, 筆記再生率を独立変数とした, 2(アクティビティ: Mediation活動なし, Mediationあり)×2(テキストタイプ: 物語文, 説明文)の分散分析(ANOVA)を行った。その後の交互作用の検定と多重比較では, 有意水準の調整方法としてHolm法を用いた。本研究では, Mediation活動が英文理解にどれくらい効果的であるかを解釈するために, 効果量dも参照した($d < 0.2$:効果量小, $d = 0.5$:効果量中, $d > 0.8$:効果量大)。

3.4.2 Mediation活動におけるやり取りの質的分析

生徒がMediation活動において, 内容理解を深めるために, どのような思考力・判断力・表現力を発揮していたのかを質的に分析した。具体的には, はじめに生徒のやり取り(発言)を文字起こしした。その後, ①思考力・判断力, および②表現力を働かせているケースを抽出した。

4 結果

4.1 筆記再生

表3にMediation活動の有無による各テキストの筆記再生結果を示す。Mediation活動後は, 物語文, 説明文とも筆記再生テストの数値が高くなった。この結果は, Mediation活動の主効果が有意であることからも支持される, $F(1, 18) = 19.01$, $p < .001$, $\eta_G^2 = .14$。また, Mediation活動の効果は読み物教材のジャンルで異なることはなかった。

■表3: 筆記再生の結果

テキスト	Mediation活動なし			Mediation活動あり			*d*
	M	95%CI	*SD*	*M*	95%CI	*SD*	
物語文	19.53	[14.50, 24.55]	11.18	30.00	[25.74, 34.26]	9.73	1.00
説明文	21.85	[17.72, 25.98]	9.42	27.05	[22.94, 31.16]	9.38	0.55

4.2 Mediation活動における思考力・判断力・表現力

Mediation活動中，話し手となった生徒たちは読んだ内容について伝える一方で，聞き手の役割を担った生徒たちも「話の内容について質問すること」を積極的に行っていた。質問の特徴として，話の中心となる人物や話題に関することが多く見られた（例1・例2の下線部参照）。特に，Mediation活動が始まった早い段階で話の全体像やキーワードを要求する聞き手からの発言（例1の波線部参照）も見られた。さらに，話し手は聞き手からの要求がなくとも，話の概要やキーワードをまず初めに伝え，詳細を伝えていた（例2の波線部参照）。これらの結果は，生徒たちが英文中の情報の重要度を考慮し，どのように相手に伝えれば良いかを適切に判断していたことを示している。

例1: Wegner（物語文）の内容

01 S1: どんな話なんですか？
02 S3: アルフレッド・ウェグナーの…
03 S2: あ，聞いたことある。
04 S1: 聞いたことある。
05 S2: え，何人ですか？
06 S3: Germanyってどこの国？
07 S4: ドイツ。
08 S2: ドイツ。男の人ですか？
09 S3: 男の人。
10 S2: まあそのアルフレッド・ウェグナーさん，何者ですか？
11 S3: まあめっちゃ話を飛ばしたら，地球の全部の陸地が昔は繋がってるみたいな…
12 S2: あー，はいはいはい。知ってる。
13 S1: あー，はいはいはい。あの人や，丹羽先生の人ですよね。

S1とS2は聞き手（質問する側），S3とS4は話し手としてMediation活動を実施

例2：Left-Handed（説明文）の内容

14 S1: えっと，世の中に生きてる人で，右利きと左利きの人がいて，世の中には両利きの人がおるっていう話で，でも，なんでほとんどの人が使われへんのかっていうのを調べて，でも…それは…。
15 S2: なんか，その専門家の人たちは，その，とりあえずなんでどっちかしか使えないかっていうのはわからんかって。とりあえず，絶対昔から自然現象で，どっちかしか使えないになってるみたいな雰囲気はわかってたらしい。だからそれで…。
16 S3: 生まれつきっていうこと？
17 S2: 生まれつきどっちかしか使えない，みたいな。

S1とS2は話し手，S3は聞き手（質問する側）としてMediation活動を実施

さらに，生徒独自の推論を働かせた質問もMediation活動のやり取りでは行われていた。例3では，Wegenerが大陸移動説を考えるに至った原因を特定しようとする質問であったり（下線部参照），物語の結末（結果）を確認するような質問が行われていた（波線部参照）。また，例4のように，話し手自身が「原因」にまつわる推論を自ら行い，話の内容をより細かく伝えているケースも見られた（下線部参照）。

例3：Wegner（物語文）の内容

18 S1: これなんで，1つの大陸からできてると思い始めたんですか？
19 S3: えっと，その人が，えっと，アフリカと南アメリカの陸がなんとなく繋がる形になってるっていうのを見て…
20 S2: その最終的に，アルフレッド・ウェグナーさんの考え方みんな認めたんでしょうか。
21 S3: えっと，死んでから。
22 S2: あー，死んでからか。あー，それは惜しいことをした。それはその認められたきっかけというか，最初はだって認められてなかったのに，何をきっかけに認められたというか，本を出してからでしょうか。
23 S4: えっと，その人が死んで，それについてもっと詳しく知りたいっていう人たちが勉強して，調べた結果，1年に2㎝ずつだんだん陸が離れていってることを実験して，本当だったんだってなって…

S3とS4は話し手，S1とS2は聞き手（質問する側）としてMediation活動を実施

例4：Left-Handed（説明文）の内容

24 S2: まあこれはまあ道具を共有，シェアしたりとかまあいっぱいあるけど，その中で，あ，ちゃうわ，あ，ほんで，じゃあなんで左利きの人が出てきたんやっていう，それは，このひとつの…
25 S1: あの，左利きやったら…
26 S2: 大きなアドバンテージができるから，そのアドバンテージができる理由としたら，例えばこの右利きの人やったら，いっつも右利きで慣れてるけど，左利きの人が来たら，その…
27 S1: 右利き。
28 S2: いつもと慣れていない，そうそう，予想しないような攻撃がくるから，アドバンテージになるよねっていう。

S1とS2は話し手としてMediation活動を実施

そのほかにMediation活動のやり取りで見られた特徴として，繰り返し（例5の下線部参照）による返答や質問が用いられていること，また聞き手からは質問以外の発言（例5の波線部参照）も多く見られた。さらに，内容理解の訂正（例6の下線部参照）も，Mediation活動を通して行われていた。また，日本語でその語彙が表現できない（Cataract：白内障）時は，その上位語（目の病気）に置き換えて，Mediation活動を継続していた（例7の下線部参照）。

例5：Eyesight（説明文）の内容

29 S4: 動物の白内障を治療するための何か。
30 S1: それってどの動物でもいいんですか。
31 S4: えっとね。はい。

32 S2: えっ, すごーい。
33 S1: えっと, 白内障になりやすい動物とか。
34 S4: 書いてないです。すいません。
35 S3: 何かを中国行く時, 送った…
36 S2: 何？
37 S3: その何かその, 多分目のレンズを中国の, まあ熊に…
38 S2: 熊？どんな熊とかも書いてないん？
39 S4: 白内障になった熊。
40 S2: あ, 白内障の熊。
41 S4: 治療しないといけない熊のためのやつを。

S1とS2は聞き手（質問する側）, S3とS4は話し手としてMediation活動を実施

例6： Left-Handed（説明文）の内容

42 S3: 左利きの割合は10%なんやろ？
43 S2: そう。世界で見たら。
44 S3: 両手利きは？
45 S2: 両手利きは, 書いてないですね。すいません。
46 S1: ごくわずかな人。
47 S2: そう, ごくわずかな人は, 両利きの人もいるよねって。

S1とS2は話し手, S3は聞き手（質問する側）としてMediation活動を実施

例7： Eyesight（説明文）の内容

48 S1: この話はなんか, コントラックス？っていうなんか最初にこんな言葉が出てきてんけど。
これは自分の目の, 目の話やねんけど, これ, 目の病気になって, 目の…
49 S2: 目のレンズや。

S1とS2は話し手としてMediation活動を実施

最後に, Mediation活動中にどのような知識・技能（観点1）, 思考力・判断力・表現力（観点2）, 学びに向かう力（観点3）を発揮したのかについて, 生徒自身のふり返りを報告する。まず, 観点1については, Mediation活動を行うことで協働的な学習による知識の確認が行われていたことがわかった。また, 相手に伝えるという活動目的が, 理解する技能につながっていたと生徒たちは感じていた。

観点1	＜知識＞ ・自分が見過ごしていたり, 意味がわからなかったりしたところを, みんなで共有することができて, 話をよく理解することができた。 ＜技能＞ ・あまり読むことのない長い説明文を読むことで, 読解力が身についたと思う。 ・人に説明することで, 自分の話への理解度が上がった。 ・相手のグループに話すために覚えることで, 記憶力が良くなることが凄く良かったです。瞬時に英文を日本語訳することで, 翻訳する力が身につくことが役立ちました。

観点1	・回を重ねるごとに記憶力が上がったと思います。翻訳をする力もついていって，しっかり読もうと思いました。 ・コミュニケーション能力が上がったこと。 ・一回やると内容がほとんど理解できたこと。

観点2では，話し手がMediation活動中に聞き手に内容を伝えながら，同時に内容の整理も行っていること，また同じチームの自分のペアの話も聞くことで，自分の理解が正しいかを確認し，より内容理解を深めていた。さらに話し手は，考えたことや理解したことを自分の言葉で伝えられていることを実感していた。

観点2	＜思考力・判断力＞ ・相手に質問されるから，その質問に答えられなかった所が，自分が理解できていなかったとわかる点。同じ内容の人がもう一人いるから，その人と言っている内容が違ったら，自分やもう一人の間違いに気づける。 ・読み解くだけでなく，その内容を伝えるので，いかに理解して，要約して，伝える事は社会に出たら何かしらで使うかもしれないと思う。 ・話をしながら，自分でも話の内容を整理できた。 ・自分がわからなかったところを友達がわかっていて理解が深まった。 ・いろんな人の同じところでも見方や意見が異なると思うので，自分の意見だったり，相手の意見だったり知れてもっと深く考えてみようと思った。 ＜表現力＞ ・人に説明する力がついたと思います。 ・出来れば英語で話した方が英語のアウトプット力が向上すると思うけど，日本語で伝えることで，英語で単語がわからない，難しいなどで言えなかった部分を補うことができた。 ・読むだけでなく説明したり，質問されたりするから，余計話の内容を本当に理解できておかないといけないけれど，自分の言葉に直して言えたらより嬉しくて楽しかった。 ・実際に他人に教えるような話し方で伝えることができて，より内容を理解することができた。

観点3では，読解活動後にMediation活動を実施することで，お互いに補いながら内容理解が深められていることを肯定的に受け止めていた。

観点3	・同じグループのもう一人の人と内容を確認した時，一緒だったら正しく読めてたんだと自信が持てた。 ・今まで知らなかった人の話などを，英語を通じて知ることができた。 ・自分の読んだ文章の内容を頑張って伝えようと思えたし，ジェスチャーも使いながら話すことができた。 ・自分の読んだもので，私は分かったところがペアの子は分かってなかったり，私が分からなかったところはペアの子が補ってくれたので，とてもよかったです。 ・互いの考えを共有できて楽しい。

観点3	・内容を自分で復習することにつながり，相手の話していることを理解しようと聞けた。 ・知らない単語を友達に聞けてすっきりする。 ・自分や友達の，伝えようと変な言葉遣いになってしまうのも面白かった。

一方で，Mediation活動においてこれらの観点に難しさを感じていることもわかった。具体的には，観点1では知識や技能が十分でないことによりMediation活動中に困難を感じていることや，ペアで理解している内容が異なっていることが挙げられる。観点2では聞いた話の内容から質問を作成することや，話し手も内容理解が不十分で細かい内容を含め質問に十分に答えられないこと，理解が進まない部分があったことが伺える。

観点1	＜知識＞ ・わからない単語があると，文章を翻訳することが難しかったです。意味を理解していないと，内容を覚えるのがよりできなくなることが分かりました。 ・伝えたくても，その意味のフレーズや単語が思いつかなかったりしてすごくもどかしかった。 ・わからない単語ばかりでカットして伝えるとそこを質問されるから，わからなかったとばかり言ってしまっていた。（質問にほぼ答えられなかった） ＜技能＞ ・英文を読んで日本語に訳しながら要約することが難しかったです。 ・文章自体が難しかったので理解するのが大変だった。 ・ペアの子との話の理解が違っていて，食い違った。 ・みんながどこの話をしているのか，わからなくなる。

観点2	＜思考力･判断力＞ ・質問されるから詳細まできっちり読み取れないといけないこと。 ・話を思い出すこと。 ・物語の内容を思い出すこと。 ・内容を忘れてしまうこと。 ・何を聞けばいいかが難しかった。 ・相手の人にめっちゃ質問されたら，理解していない部分の時は難しかった。 ・相手がしっかりと理解しきっていないこと。 ・質問を考えるのも難しかった。 ・質問に答える際，具体的な情報，（時間･人名･人数など）を入れて説明するのが難しかった。 ・人の名前などを覚えられていないときがあって，伝えにくいときがあった。 ・相手の質問に対して，しっかりと要約して返答することが難しかったです。 ・話の読みにもれがあったら，質問に答えられないこと。 ＜表現力＞ ・分かりやすく伝えること。 ・話の内容をしっかりと理解していないといけなかったので，上手く説明できないこともあった。 ・ペアで相手のペアに話の内容を話すときに話の内容が難しかったため，分からなかったところを無理やり説明するのが難しかったので，日本語で相手に伝えることがとても難しかった。 ・時によくわからないことがあり，それを相手に伝えることができなかった。

5 考察

5.1 内容理解の向上のついて

表3にあるように，Mediation活動を行うことで物語文と説明文の両方のジャンルで生徒たちの字義的理解が向上し，内容理解が促された。生徒たちのMediation活動のやり取りからは，開始直後から話の概要について触れるなど，話の全体像をまず把握し，そこからさらに内容理解に必要な個々の情報を選んで質問するなどを行っていた。書かれていた内容をそのまま話すよりも，全体像や中心となる人物・内容について話をする方が，内容理解を促進させるためには必要と，生徒自ら考え判断した結果と言える。また「原因」や「結果」に関する推論質問が行われていたことも，話の内容理解が促された一因と言える。加えて興味深いことは，話し手となった生徒が，聞き手が疑問に思うであろうことを想定し，自らの判断で推論した内容（原因）を伝えていたことである。相手を慮り，得られた情報を必要に応じて伝えるというこの一連の流れは，生徒たちの「思考力・判断力・表現力」を大いに刺激していると言える。

またMediation活動では，繰り返しが頻繁に見られた。繰り返しは，話し手が少ない労力で追加情報を伝えることができ，すでに議論されたことに新しい内容を加えることにつながる（Tannen，1987）。Tannen（1987）は，繰り返しをproduction，comprehension，connection，interaction，coherenceの5つの目的別に分類した。Mediation活動では話し手／聞き手の双方が繰り返しを行っていたが，お互いが前述のつながりを意識し，話の内容を一貫して理解しようとする目的で行われたと考えられる。また聞き手からの繰り返しを伴った質問には，理解したことの確認とさらなる追加説明の要求を暗に意味している。そのため，話し手は更なる説明を行う必要があり，より一貫して話を理解することにつながったと考えられる。またMediation活動は，聞き手からの質問以外の発言も注目に値する。三宮・山口（2019）は聞き手からの肯定的なあいづちが，動機付けや発想量を高めやすいことを明らかにした。これは，聞き手からのあいづちにより話し手は，聞き手が関心を持っているという安心感を得られ，アイデア生成が活発となる。Mediation活動内での聞き手のあいづちや感想は，無意識であっても話し手の話す意欲の向上につながっており，さらなる情報の追加を促し，内容理解を促進させるものと考えられる。

Kai（2008）は，生徒の熟達度のレベルが読解の理解度に影響があることを述べた。Mediation活動の中にも，生徒が語彙を理解できていないことがわかる発言が見られる。一方で例7には“cataracts（白内障）”にあたる言葉は出てこない。しかし生徒たちは「白内障」という言葉がわからなくても，「目の病気」という上位語の言葉に言い換えてMediation活動を展開させていた。生徒たちが自らの言葉で伝えようと取り組んだ結果と言える。

Mediation活動は，能力差の影響からペアを組んだチーム内でも話の理解度に差があることや，一方が間違えて理解をしていたことも生徒の感想からも伺える。しかしMediation活動を通して，チーム内での理解度の差や解釈の違いを埋め，そして間違いがあれば修正する様子が見られた。Mediation活動には，生徒の思考力・判断力・表現力を試される場が多々ある。内容理解を深めるための過程は様々あり，行われるやり取りがグループ内の一貫した理解の形成に大きく関わっていると言える。

5.2 生徒によるMediation活動のふりかえり

自由記述の内容から，話の内容を説明することで，内容を整理することができたとふりかえっている生徒がいた。使用する言語についても母語（日本語）が有効であったことに言及している生徒もおり，母語で行う方が生徒の言語的な負担を軽減できること，そして生徒が自分の言葉で語ることの面白さを感じていた。Mediation活動を通して，話し手の役割を担った生徒が，答えられえない質問やうまく説明できない場面に直面することで，自身の内容理解が十分でないと認識していた。このことからMediation活動には，相

手にとって必要な情報を決定し，わかりやすく伝えようとする思考力や判断力を育てる一方で，自身の内容理解を促すために自分が必要とする情報が何なのか，ということを考えるきっかけを生徒に与えていることがわかった。他者の視点が呼び水となり，生徒たちは内容理解をさらに深めようとする意識を高めることにもつながる。

しかし一方で，活動中に語彙や文法の知識不足から母語であっても上手く伝えられなかった生徒がいたことも，生徒の感想から確認できる。また話を聞いて内容に関する質問をすぐに行うことの難しさを感じている生徒もいた。しかし知らない語を，関連する語で言いかえて説明するといった工夫が見られていることから，生徒自身も気づかないところで創意工夫し，思考を巡らせながら，Mediation活動を行っていたと言える。そのような生徒自身が相手に伝えるために工夫していたことを生徒に伝えることも，生徒たちの意欲を高め，それがMediation活動をさらに充実させることにつながる。

6 結論

6.1 本研究のふりかえり

本実践の英文読解後のMediation活動なし・ありの比較において，Mediation活動を行うことで物語文・説明文の両方の英文理解が促されていた（RQ1）。それはMediation活動には，生徒の思考力・判断力・表現力が必要となる場面が数多くあることが挙げられる。さらに，RQ2に対して，Mediation活動でのやり取りから，話し手となる生徒は，早い段階で話の概要やキーワードについて語ろうとすること，また聞き手も話の概要をまずは聞こうとする傾向があった。そしてそこからキーワードとなる内容や人物のことを深く掘り下げていた。話し手の生徒は「相手にわかりやすく伝える」という目的の中，「相手は何を求めているのか」，「どのように伝えるべきか」など，内容を整理しながら，状況や場面から必要なことを考え判断し，その内容を伝えていた。聞き手はキーワードとなる語や内容を話し手の話の中から判断し，推論質問を含め，内容理解のために必要な質問を行っていた。日常を反映した言語使用というMediationの特徴のもと，やり取りの中で行われている繰り返しやあいづちもコミュニケーションを促し，Mediation活動でのやり取りがさらに活発になり，よって内容理解が深められることにつながったと考えられる。また話し手は聞き手からの質問に答えることで，自らの内容理解の程度を確認していること，そして内容理解を深めるためにどのような情報が必要となるのかを考えていた。文章中のわからない語については別の言葉に言い換えるなど，Mediation活動では生徒たちが自らの言葉で語る工夫を行っている様子が見られた。

6.2 今後に向けて

Mediation活動は生徒の英文読解活動に効果的であることが確認され，かつ生徒たちも好意的に受け止めていたことが伺える。一方で，活動を行う生徒たちは，聞いた話に対してその場で質問をすることに難しさを感じていた。また教師にとってもクラスサイズが大きくなるほど生徒のレベルにあった教材を選定することが難しくなることが課題として挙げられる。さらに今回のMediation活動を実際の授業の中にどのように取り入れていくことができるのかということも今後の考えていくべき課題である。中学校・高等学校の現場では，英文読解の際に教科書を使用することがほとんどであるため，教科書を使用したMediation活動の実施などを行っていきたい。また語彙や文法の理解が原因となってMediationを十分に行えない生徒たちへの指示の出し方や，絵やキーワードが書いてあるカードを使用させることなどでの支援の方法を考える必要がある。加えて，ペアやグループの組ませ方でもMediationの中身や理解度が変わる可能性があることから，それらも含め更なる検証が必要である。

読解活動を一人で行うことには限界がある。だからこそMediation活動を通して他者からの視点を新

たに加えて読み物教材に向き合うことは，とても有意義なことと言える。Mediation活動を通して伝え合うことの楽しさを感じ取りつつ，生徒たちが自らの力で読み物教材への理解を深められるよう今後も研究を重ねたい。

謝辞

今回本研究を実施・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 日本英語検定協会の皆様，選考委員の先生方に深くお礼申し上げます。とりわけ，関西大学の竹内理先生には報告書作成に至るまで温かいご助言・ご指導を賜りましたことに厚く御礼申し上げます。そして，神戸市外国語大学の濱田彰先生には本研究の構想段階から細やかなご指導をいただきました。この場をお借りして御礼申し上げます。最後に本研究の実施にあたり協力してくださった先生，そして一緒に取り組んでくれた生徒の皆さんに心より感謝します。ありがとうございました。

引用文献

Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment (Companion volume).* Council of Europe. http://www.coe.int/lang-cefr

Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E., & Seipel, B. (2020). Inferial comprehension differences between narrative and expository texts: a systematic review and meta-analysis. *Reading and Writing, 33,* 2223-2248. https://doi.org/10.1007/s11145-020-100442

Coste, D., & Cavalli, M. (2015). Education, mobility, otherness -the mediation functions of school. Council of Europe. https://rm.coe.int/16807367ee

Gibson, A., Gold, J., & Sgouros, C. (2003). The power of story retelling. *The tutor,* 1-11.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* Cambridge University Press.

Grabe, W., & Stoller, F. L. (2020). *Teaching and researching reading* (3rd ed.). Routledge.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 101(3), 371-395. https://doi.org/10.1037/0033-295X.101.3.371

Hamada, A., & Takaki, S. (2021a). Approximate replication of Matsuda and Gobel (2004) for psychometric validation of Foreign Language Reading Anxiety Scale. *Language Teaching, 54(*4), 535-551. https://doi.org/10.1017/S0261444819000296

Hamada, A., & Takaki, S. (2021b). Effects of multidimensional foreign language reading anxiety on achievement in Japanese EFL classrooms. *System, 101,* 102613. https://doi.org/10.10 16/j.system.2021.102613

Horiba, Y. (2000). Reader control in reading: Effects of language competence, text type, and task. *Discourse Processes, 29*(3), 223-267. https://doi.org/10.1207/s15326950dp2903_3

池田周 (1994). 読解における推論タイプとその役割について，『中国地区英語教育学会研究紀要』, No.24, 97-106

甲斐あかり(2008). 英文読解テストとしての再話課題の有効性の検証：―テキストタイプ，産出言語，採点方法の妥当性を中心として―，『STEP Bulletin』, 20, 76-94.

Kai, A. (2008). The effects of retelling on narrative comprehension: Focusing on learner's L2 proficiency and the importance of text information. ARELE: Annual *Review of English Language Education in Japan,* 19, 21-30. https://doi.org/10.20581/arele.19.0_21

Kai, A. (2009). Achieving global coherence through retelling. ARELE: *Annual Review of English Language Education in Japan, 20,* 41-50. https://doi.org/10.20581/arele.20.0_41

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294-303. https://doi.org/10.1037/0003-066X.49.4.294

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach.* Cambridge University Press.

Koskinen, P. S., Gambrell, L. B., Kapinus, B. A., & Hearthington, B. S. (1988). Retelling: A strategy for enhancing students' reading comprehension. *The Reading Teacher, 41*(9), 892-896.

文部科学省 (2017).『中学校学習指導要領解説外国語編』, 文部科学省.

長沼君主 (2018).【新企画リレーコラム】言語能力育成を考える 第3回「CEFR から読み解くタスクベースの言語能力発達」前編 (arcle.jp) https://www. arcle.jp/note/2018/0028.html

長沼君主 (2018).【新企画リレーコラム】言語能力育成を考える 第3回「CEFR から読み解くタスクベースの言語能力発達」後編 (arcle.jp) https://www. arcle.jp/note/2018/0029.html

North, B., & Piccardo, E. (2016). Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR): A Council of Europe project. *Language Teaching, 49*(3), 455-459. https://doi.org/10.1017/S0261444816000100

Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the one proximal development

引用文献

and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51-78). Oxford University Press.

Ray, M. N., & Meyer, B. J. (2011). Individual differences in children's knowledge of expository text structures: A review of literature. *International Electronic Journal of Elementary Education, 4*(1), 67-82.

三宮真智子・山口洋介. (2019). 発想に及ぼすあいづちの種類の効果.『心理学研究』, 90(3), 301-307.

Schallert, D. L., & Tierney, R. J. (1980). Learning from Expository Text: The Interaction of Text Structure with Reader Characteristics.

Singer, M., Andruslak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & cognition, 20*(5), 539-548.

Stathopoulou, M. (2015). *Cross-language mediation in foreign language teaching and testing.* Multilingual Matters.

Tannen, D. (1987). Repetition in Conversation: Toward a Poetic of Talk. *Languages,* 574-605.

Thorne, S. L. (2000). Second language acquisition theory and the truth. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 219-243). Oxford University Press.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive psychology, 9*(1), 77-110.

投野由紀夫・根岸雅史(2020).『教材・テストCEFR-Jリソースブック』, 東京:大修館書店

卯城祐司(2009).『英語リーディング科学―「読めたつもり」の謎を解く』, 東京:研究社

Ushiro, Y., Hamada, A., Kimura, Y., Nahatame, S., Hosoda, M., Kato, D., & Watanabe, Y. (2015). Building causally coherent mental representations and learning from expository texts among Japanese EFL readers. *JACET Journal, 59,* 131-150.

Ushiro, Y., Hamada, A., Mori, Y., Hosoda, M., Tada, G., Kamimura, K., & Okawara, N. (2018). Goal-oriented L2 reading processes in maintaining the coherence of narrative comprehension. *JACET Journal, 62,* 109-128. https://doi.org/10.32234/jacetjournal.62.0_109

van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*(8), 1081-1087. https://doi.org/10.3758/bf03206376

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

第35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

AIツールを活用して日本人中高生の学習の主体性を高めることをねらったライティング指導

研究者:松井　市子　新潟県／新潟県立津南中等教育学校　教諭
《研究助言者:斉田 智里》

概要

本研究は,「主体性」を「自律学習能力」と捉え, 日本人中高生を対象にAIツールを活用して学習の主体性を高めることをねらったライティングタスクを指導に取り入れた。生徒は, タスクでフィードバックの特徴が異なる複数のAIツールから主体的に選び, 教師の支援がほとんどなくてもAIが出した結果(フィードバック)を読み取り, 自力でのCEFRレベルアップをねらった。アンケートを用いた分析の結果, CEFR_A学習者は自力でのレベルアップは難しく, 教師不在への不安も大きいが, 見通しを立てたり振り返ったりして自らの学習を自覚的に捉えるなど自律学習能力の育成が可能であることが分かった。学年の特性を踏まえて本タスクを取り入れることで, 教師のフィードバックの負担減や学習者のライティング活動機会増が期待でき, 教育AIを活用して公立学校で「1人1台端末」を使いながら最適な個別学習のための持続可能なライティングタスクを生徒主体で推進できる可能性が示された。

1 はじめに

1.1 GIGAスクール構想における教育観

GIGAスクール構想の下「1人1台端末」の学習環境が整備され, 教育観のパラダイムシフトの必要性が高まっている(赤堀, 2022)。高等学校学習指導要領外国語編英語編(2018)は, 学習が進むにつれて教師の支援がほとんどなくても学習を成立させる資質や能力の育成を求めており, これまでの教師主導から生徒主体の授業観に基づいた授業実践が迫られている。また, 文部科学省(2019)は, 誰一人取り残すことのない「個別最適化された学び」の実現のために, 支援ツールとしてのICT環境と先端技術が不可欠であると示している。森本(2020)は, 個別最適化された学びを実現させるには, 教育AIと学習記録データの活用が有効であると指摘している。このような背景を踏まえると, 外国語教育において教育AIの助けを借りながら生徒主体で個別最適化した学習を「1人1台端末」を有効に活用して推進するという教育観が必要である。

1.2 教育AI

教育AIに関しては様々な目的や用途で開発されているが(Zhai, et al, 2021), 張(2019)は教育における「AIの利活用(AI in education)」から「教育AI(Educational AI)」の段階に入ったとして, アメリカの

学校教育を例に「チュートリアル」「個人学習」「テスト」「自動化タスク」としての教育AIを示している。日本の外国語教育では依然として「書くこと」や「話すこと」のアウトプット能力の育成が課題とされており（文部科学省，2018），「日本の40人学級における指導者への過重負荷やタスク開発等に関わる知識技能不足等，技能伸長の阻害要因」（Koizumi, et al, 2022）を考えると，アウトプット能力を育成するための教育AIの活用が期待される。松井（2020）は，AIによるライティング自動添削ツールであるWrite & Improve（W & I）（https://writeandimprove.com）を使った先行研究で，CEFR_A2学習者のライティング力を育成するのにAIツールが効果的であるという結論を示した。また，ライティング力の伸びが見られなかったCEFR_A1～B1学習者にはAIへの不信感があることを示した。これはAIリテラシー教育に関わる分野で，「個人学習」を最適化するためには学習者が「人工知能を適切に活用するための思考力・判断力を働かせながら，AIが出した結果に対してその求め方（経過）を読み取り，技術の長所と短所を踏まえた上で，活用について考える」（向田，2022）必要がある。

1.3 主体性の育成

平成29～30年改訂学習指導要領では，評価の観点に「学びに向かう力（主体性）」が加わった。主体性は，①「知識及び技能を獲得したり，思考力，判断力，表現力等を身に付けたりすることに向けた粘り強い取組を行おうとしている側面」，②「①の粘り強い取組を行う中で，自らの学習を調整しようとする側面」，③「①②の側面と併せて，言語活動への取組に関して見通しを立てたり振り返ったりして自らの学習を自覚的に捉えている状況についても，特定の領域・単元だけではなく，年間を通して評価する」と示している（国立教育政策研究所，2021）。

外国語教育では，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）が習熟度を示す指標として活用されている。大木（2021）は，学校教育で外国語力を高めるために「自律学習能力」の育成が重要だとして，CEFR考案者の1人であるHenri Holecが提唱した「自律学習能力」に関する質問項目を用いた学習者の自己評価による内省を促している。表1は，その質問項目を示したもので，本研究に合わせて一部加工した（下線部分）。

■表1:「自律学習能力」に関する質問項目

1. 目標設定
1.1 私には私なりの外国語学習の目標がある。
2. 学習するべき内容の明確化
2.1 私は外国語学習の目標を達成するために何をしたら良いのか分かっている。
2.2 私は授業時間外に何を勉強したら良いのか分かっている。
2.3 私は何が不得意なのか分かっている。
2.4 授業時間外に外国語を勉強するときには，自分の不得意なところに特に時間を割いている。
2.5 私には何が重要なのか，また何に力を入れて学習すればよいのか分かっている。
3. 学習方法・教材の明確化
3.1 私はどのような方法を用いて外国語を学習すればよいのか分かっている。
3.2 外国語を勉強していて，分からないことがあった場合，私はどのようにして解決したら良いのか知っている。
3.3 私はどのような教材を用いて勉強すれば良いのか分かっている。
3.4 私は私の持っている母語の知識を外国語を学習するときに用いている。
3.5 私は私の持っている母語の知識が外国語の学習の時に役立つことを知っている。

4. 学習のコントロール・自己管理
4.1 授業時間外に（教師のいないところで），外国語の学習をするのは不安だ。
4.2 教師がいなくても，適切な教材があれば，私は外国語はマスターできると思う。
4.3 教師の説明がなくても，適切な教材があれば，私は外国語を理解することができる。
4.4 私は外国語を学習するための時間をどのようにして確保したらよいのか分かっている。
4.5 私は外国語を学習するのに必要な時間を授業時間外でも十分に確保している。
5. 評価
5.1 私は今学習していることを理解しているかどうか自分で確かめることができる。
5.2 私は私自身の外国語の実力を自分で評価する方法を知っている。
5.3 私は時々私自身の外国語の実力を自分で評価している。

自律学習能力を育成するには，それを構成する「1. 目標設定」「2. 学習すべき内容の明確化」「3. 学習方法・教材の明確化」「4. 学習のコントロール・自己管理」「5. 評価」を理解して実施する学習者を養成し，教師から受ける援助の程度を徐々に少なくして学習の決定を学習者に委ねていく必要がある（大木，2021，第3章）。これは文部科学省（2018）の「支援がほとんどなくても学習を成立させる」という高校の最終的段階で期待される姿と重なる部分である。また，「自律性」は教育心理学の自己決定理論や学習者エンゲージメント研究で「主体性」と関連付けて説明され（西田，2022; マーサー＆ドルニェイ，2022），学習者が主体性（ownership, agency）を持って取り組むには有能性（やればできるという期待感や達成感）や自律性（学習に自分で責任を負っているという感覚），自己統制感や粘り強さの育成が必要で，教師はコーチとして学習者に学習の責任を置き，学習の進歩を可視化する手だてを示す役割が考えられる。

国立教育政策研究所（2021）は，「技能」を「主体性」と一体的に評価する方法を示しているが，自律学習能力を主体性と一体的に評価する方法も可能だ。特に，生徒が粘り強い取組を通して自らの学習を調整しようと言語活動への取組に見通しを立てたり振り返ったりして自覚的に学習を捉えて主体性を高めるためには，教師の役割に関する教育観のパラダイムシフトが求められる。生徒が外国語学習の目標を主体的に設定し，学習すべき内容や方法，教材を明確化し，学習をコントロール・自己管理して，評価を下す姿は目指すべき主体的学習の側面であり，主体性が高い学習を展開できるかどうかの判断材料になりえる。

2 研究内容

2.1 研究の目的

本研究の目的は、日本人中高生がAIツールを活用してライティングタスクに取り組んだ時に，主体性が高い学習を展開できるかどうかを明らかにすることである。本研究の特色は，間接的フィードバックを与えるW & I以外にも直接的フィードバックを与えるGrammarly（https://app.grammarly.com）や語彙レベルのフィードバックを与えるCEFR-based Vocabulary Level Analyzer ver.2（CVLA ver.2）（https://cvla.langebu.jp, Uchida & Negishi, 2018）も併用することである。松井（2020）はW & Iを使ったライティングタスクにおいてCEFR_A2学習者のレベルアップが見られたこと，また，AIツールへの不信感がCEFR_A1～B1学習者のレベルアップの阻害要因である可能性を示した。本研究では，フィードバックの特徴が異なる複数の選択肢から生徒主体でAIツールを選び，思考力や判断力を働かせて，教師の支援がほとんどなくてもAIが出した結果（フィードバック）を読み取り，個別学習に活用していくことをねらった。また，AIツールを活用することで教師（人間）のフィードバックの負担を減らしたり，瞬時のフィードバッ

クにより学習者の学習機会を豊富に確保したりすることをねらった。研究で使用した3つのAIツールは，英文インプットに対して，CEFRレベル判定や文法ミスの訂正提案などのフィードバックを与える特化型AI（narrow AI）で，データ処理に特化したツールである。無料のインターネットサービスで，生徒が自分のライティングのCEFRレベルを知ったり，文法ミスに気付けたりするため，個別最適に「1人1台端末」環境を有効に活用できるツールだと期待して，この3つを選択した。また，本原稿の執筆中にも生成AIの教育現場での使用に関する暫定的なガイドライン（文部科学省，2023）が示されたが，本研究で使用した3つのAIツールは，データを生成する生成AIとは異なる認識である。

本研究のリサーチクエスチョン（RQ）は以下の通りである：

RQ1	AIツールを使って自力でライティングのCEFRレベルをアップできるか
RQ2	AIは教師（人間）より「書く意欲」や「書く技能」を高めると学習者に思われるか
RQ3	学習者のCEFRレベルで選ぶAIツールに差があるか
RQ4	学習者のCEFRレベルでAIフィードバックの理解度に差があるか
RQ5	学習者のCEFRレベルや学年で自律学習能力スコアに差があるか

外国語学習においてもICT（情報通信ネットワーク）を利活用して言語運用能力や言語活動を促進することが，期待されている（文部科学省，2020）。本研究では，そのようなICT活用環境において，日本人中高生がライティング力を育成するときに，フィードバックに関して教師の役割に代わる教育AIをどう利活用したのか，自律学習能力の育成を目指したライティング指導を通して主体性が高い学習を展開したのかを検証した。なお，RQ1～RQ2は，松井（2020）の結果と比較して検証した。

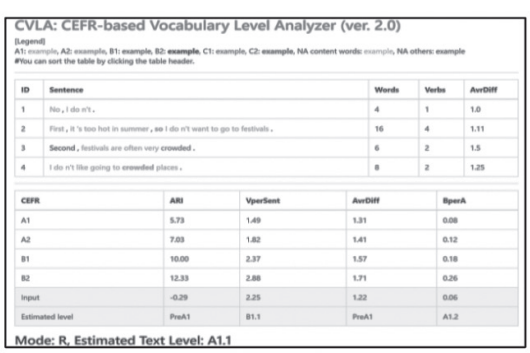
2.2 研究の対象

本研究の筆者の前任校である公立中等教育学校1～6年334名を対象に，2022年度の中学校「英語」および高等学校「コミュニケーション英語」の48分授業6～7回を使って授業実践をした。授業実践者は著者を含む授業担当教師8名であった。対象生徒のうち，後期生（高校生）の英語運用能力はGTEC_Advanced_2022でCEFR_A2.1～2.2であった。

2.3 研究の手順

はじめに研究の目的や手順を英語科会でグーグルクラスルームを使用しながら筆者が説明した。タスクは「事前タスク」「練習タスク（3回分）」「事後タスク」「遅延タスク」という構成で，授業担当者と筆者で

■図1：W&Iのタスクとフィードバック例（3級）



■図2：CVLAのフィードバック例（3級）

役割や実践日を検討した後，タスクの手順（資料1）を各クラスのグーグルクラスルームに投稿して実践を行った（事前・練習・事後タスクは，3〜6年生が前期中間考査後，1,2年生が後期中間考査後，遅延タスクは後期期末考査後）。前期生（中学生）では端末の取り扱いやタスク手順において差が生じる可能性があるため，説明動画もクラスルームに投稿した。タスクのプロセスで生徒から質問が出た場合は教員が対応するが，それ以外は教員から生徒に働きかけないという共通理解を図った。

事前・練習・事後タスクは実用英語技能検定1〜3級のライティングセクション問題（資料2）を使用して，W&Iのworkbookで取り組めるようにした（図1参照）。練習タスクは英検の過去問題から選んで5つ用意し，練習タスク1〜3を授業で使用し，練習タスク4〜5は自主的に取り組みたい人用だと告げた。

事前タスクでは，生徒はまずタスクのねらいや特性について教師から説明を聞き，取り組む英検級を選択後，W ＆ Iでタスクに取り組んだ。ドキュメントに作成した英文データやAIフィードバックを記録後，アンケートに回答するという流れで実施した。このタスクのねらいは，取り組む級の英文とAIツールフィードバックの特性を理解させることで，そのためにW&IとCVLAを使わせ，各ツールでどのようなフィードバックが得られるかを比較させた（図1と図2の比較）。

練習タスクは事前タスクと間を空けず続けて3回実施した。生徒は事前タスクと同じ級を選択して，W ＆ Iで約20分間ライティングタスクに取り組んだ。Grammarlyの使用方法の説明を教師から聞き，3つのツールから好きなものを選んで，約20分間校正に取り組んだ。最後にCEFRレベルやAIフィードバック（図3）をドキュメントに記録して，アンケートに回答した。このタスクのねらいはAIツールに慣れ，AIツールの助けを借りながら自力でCEFRレベルをアップさせることであった。

■図3：Grammarlyのフィードバック例（3級）

事後タスクは練習タスクと同じ手順で，事前タスクと同じ問題を使用して，事前タスクから1カ月程度空けて実施した。このタスクのねらいはAIツールを使った学習によりライティング力を伸ばせたかどうかを自覚させることで，アンケートに回答させて振り返らせた。

タスクで使用したAIツールの特性は表2の通りである。W ＆ Iはメールアドレスでアカウントを作成して使用する。教師はworkbookを作成でき，生徒もメールアドレスでアカウントを作成することで，生徒は無料で学習記録を残せる。生徒の学習記録を教師が閲覧したいときは有料版を使用する。Grammarlyもメールアドレスでアカウントを作成して使用することで学習記録が残せる。本研究のねらいと同程度の目的で活用する場合は，無料版でも支障はない。

■表2：タスクで使用したAIツールの特性

	アカウント	使用手段	フィードバック
W ＆ I	メールアドレスで作成	ブラウザ・アプリ	間接的に語と文に関して（CEFR判定付）
CVLA	不要	ブラウザ	語と文に関して（CEFR判定付）
Grammarly	メールアドレスで作成	ブラウザ・アプリ	直接的に文法に関して

遅延タスクは事後タスクから数ヶ月～半年後のCEFRレベルを知り，自律学習能力に関して表1の「4. 学習のコントロール・自己管理」と「5. 評価」を使って，ライティングの学習でAIツールを利活用することに関する意識を把握する目的で実施した。前期生は，まとまりのあるライティングタスクを日頃実施していないため，CEFR-Jのライティングの CAN-DO タスク（http://cefr-j.org/download.html#cefrj_testasks）を利用し，自分で解答できそうだと思われるCEFRレベルを生徒自身で選択して，約15分取り組んだ。タスクはその出来具合をABCの3基準で評価し，模範解答を参照してB以上だと生徒自ら判断したタスクレベルを遅延アンケートに自己申告させた。AIツールの活用を促したが，必須ではなかった。後期生は，後期期末考査のライティングセクションの問題（50～100語程度の意見文や説明文作成問題）を使って，複数のAIツールを活用して約15分間校正し，最終的なCEFRレベルを遅延アンケートに自己申告させた。遅延タスク後，教員にもアンケートを実施した（資料3参照）。

2.4 結果の分析

実践効果はアンケートの量的データをJASP ver.0.11.1とjs-STAR_XR+を使用して分析した。また，教員アンケートの質的データを用いて実践効果との関連を分析した。

3 研究結果と考察

3.1 事前・事後タスクの英検級別人数

研究対象生徒334名のうち，事前・事後アンケート両方に回答し，かつ事前・練習・事後タスクで同じ級に取り組んだ151名を分析の対象とした（対象人数が半減した大きな理由は，タスクに没頭しすぎてアンケートの回答時間内に回答しなかったためだと考えられる）。表3はタスクで取り組んだ英検の級別の学年ごとの人数を示したものである。学年が上がるごとに取り組むタスクの級が上がり（1年 3級，2年3～準2級，3年3～2級，4年3～準1級，5年2～準1級，6年準2～準1級），取り組みたい級の幅も広がることが見取れ，生徒のニーズが多様化していることが分かった。

中学生では低学年ほどタスクに取り組む比率が低く（1年0.29，2年0.43，3年0.79），タスク難度が高かった可能性が示唆された。特に，1,2年生はアカウント管理やドキュメント等の操作に不慣れな生徒が多く，タスクに取り組み始めるまでに時間を要した。3年生ではそのような様子はほとんど見られなかった。

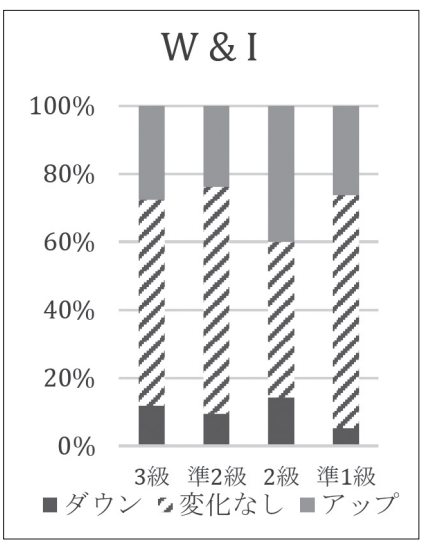
■表3: 事前・事後ライティングタスクの英検級別人数（*n*=151）

アカウント	学年						
	1	2	3	4	5	6	計
3級	23	28	23	2	0	0	76
準2級	0	4	10	1	0	6	21
2級	0	0	1	9	12	13	35
準1級	0	0	0	2	4	13	19
計	23	32	34	14	16	32	151
在籍	78	73	43	46	38	52	334
比率	0.29	0.43	0.79	0.30	0.42	0.61	0.45

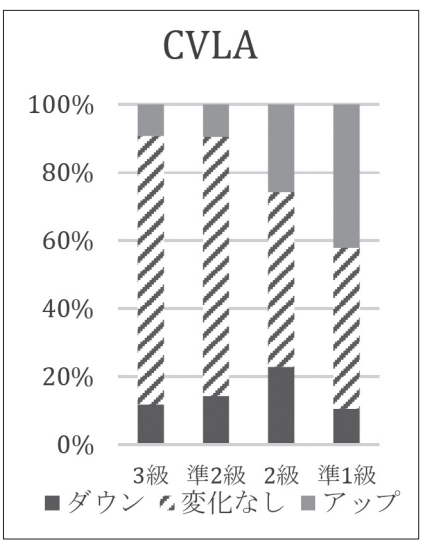
高校生では学年が上がるにつれタスクに取り組む比率は上がったものの(4年0.30, 5年0.42, 6年0.61), 取り組む級が上がりタスクに要する時間が伸びることから, 制限時間いっぱいタスクに取り組み, 結果的にアンケートに取り組まなかったことが低い比率の原因だと思われる。タスクに没頭する姿や個別最適化した多様な学びの姿が見られた。

3.1.1 RQ1:AIツールを使って自力でライティングのCEFRレベルをアップできるか

表4は事前アンケート項目5「事前タスクのW & IとCVLAのCEFRレベル」と事後アンケート項目2「事後タスクのW & IとCVLAのCEFRレベル」に関して, 事前・事後タスクで取り組んだ英検の級別に示したもので, 図4, 図5はCEFRの事前・事後の変化(【事後でダウン:表4の網掛下】【変化なし:表4の網掛】【事後でアップ:表4の網掛上】)を示したものである。英検の級ごとに実施したχ^2検定の結果, W&Iでは2級以外はばらつきが1%水準で有意であった(3級$\chi^2(2)=28.13$, 準2級$\chi^2(2)=11.14$, 準1級$\chi^2(2)=11.79$)。ライアンの名義水準を用いた多重比較の結果, 準2級, 準1級でCEFRレベルがダウンした人が少なく, 3級は変化なしの人がレベルアップやダウンした人より多かった。



■図4: W&IのCEFR変化の割合



■図5: CVLAのCEFR変化の割合

CVLAでは3級と準2級はばらつきが1%水準で有意であった(3級$\chi^2(2)=71.24$, 準2級$\chi^2(2)=17.43$)。ライアンの名義水準を用いた多重比較の結果, 3級でも準2級でも変化なしの人がレベルアップやダウンした人より多かった。まとめると, RQ1「AIツールを使って自力でライティングのCEFRレベルをアップできるか」に関して, CEFR_A学習者がAIツールを使用することでライティングのCEFRレベルが下がる可能性は低いが, 変わらない可能性も高いことが示唆された。これは松井(2020)の「CEFR_A2学習者はW&Iを使ってCEFRレベルをアップした」という結果を支持するものではなかった。

3.1.2 RQ2:AIは教師(人間)より「書く意欲」や「書く技能」を高めると学習者に思われるか

表5は, 事前・事後タスクで取り組んだ英検の級別に事後アンケート項目11「書く意欲について」および項目13「書く技能について」【どちらも高まらない(両方×)】【AIの方が高まる(AI)】【先生の方が高まる(人間)】【両方高まる(両方○)】かを示したものである(図6, 図7はその割合)。なお, 事前・事後タスクの間, 教師(人間)はライティングの添削を行わなかった。タスクで取り組んだ英検の級別に, AIと回答した人と教師(人間)と回答した人の数を直接確率計算した。「書く意欲」に関しては, どの級でも有意差は見られなかった(3級$p=0.44$, 準2級$p=0.50$, 2級$p=0.50$, 準1級$p=0.50$, 片側検定)。「書く意欲」に関しては, 2級と

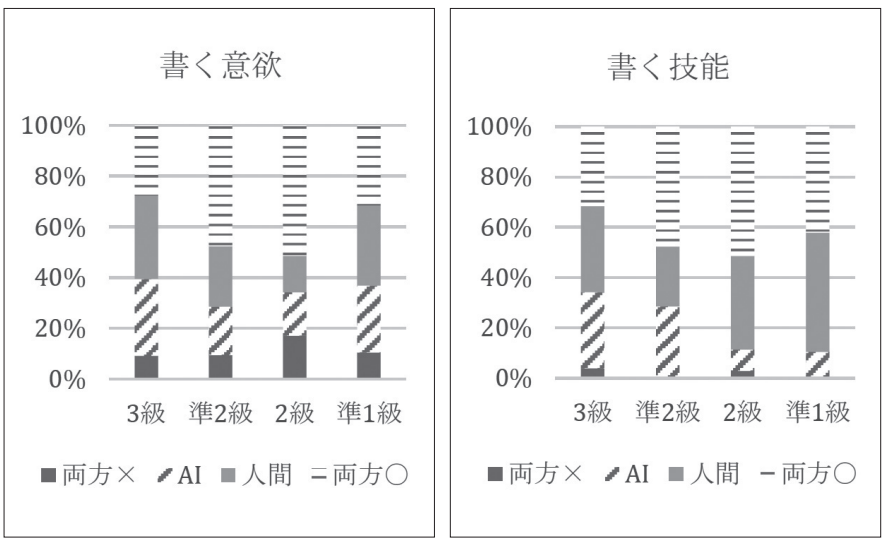
■表4: 事前・事後タスクの英検級別CEFRレベル人数 (*n*=151)

タスク		後 [W&I]					後 [CVLA]						
	前	A1	A2	B1	B2	計	A1	A2	B1	B2	C1	C2	計
3級	A1	41	20	0	0	61	55	6	1	0	0	0	62
	A2	8	5	1	0	14	5	4	0	0	0	0	9
	B1	0	0	0	0	0	2	1	0	0	0	0	3
	B2	0	1	0	0	1	0	0	0	1	0	0	1
	C1	0	0	0	0	0	1	0	0	0	0	0	1
	C2	0	0	0	0	0	0	0	0	0	0	0	0
	計	49	26	1	0	76	60	11	1	1	0	0	76
準2級	A1	3	5	0	0	8	8	2	0	0	0	0	10
	A2	1	11	0	0	12	2	6	0	0	0	0	8
	B1	0	1	0	0	1	0	1	2	0	0	0	3
	B2	0	0	0	0	0	0	0	0	0	0	0	0
	C1	0	0	0	0	0	0	0	0	0	0	0	0
	C2	0	0	0	0	0	0	0	0	0	0	0	0
	計	4	17	0	0	21	10	9	2	0	0	0	21
2級	A1	0	5	2	0	7	1	1	0	0	0	0	2
	A2	0	12	7	0	19	2	14	7	0	0	0	23
	B1	0	4	4	0	8	0	2	3	1	0	0	6
	B2	0	1	0	0	1	0	2	0	0	0	0	2
	C1	0	0	0	0	0	0	1	0	1	0	0	2
	C2	0	0	0	0	0	0	0	0	0	0	0	0
	計	0	22	13	0	35	3	20	10	2	0	0	35
準1級	A1	0	0	0	0	0	0	0	0	0	0	0	0
	A2	0	0	3	0	3	0	0	1	1	0	0	2
	B1	0	0	13	2	15	0	1	2	1	2	0	6
	B2	0	0	1	0	1	0	0	1	4	2	0	7
	C1	0	0	0	0	0	0	0	0	0	2	1	3
	C2	0	0	0	0	0	0	0	0	0	0	1	1
	計	0	0	17	2	19	0	1	4	6	6	2	19

■表5: 事前・事後タスクの英検級別「書く意欲」および「書く技能」に関するフィードバックの選択人数 (*n*=151)

タスク	書く意欲				書く技能			
	両方×	AI	人間	両方○	両方×	AI	人間	両方○
3級	7	23	25	21	3	23	26	24
準2級	2	4	5	10	0	6	5	10
2級	6	6	5	18	1	3	13	18
準1級	2	5	6	6	0	2	9	8

準1級でAIより先生（人間）の方が高まると回答した人が有意に多かった（2級p=0.01，準1級p=0.03，片側検定）。まとめると，RQ2「書く意欲を高めたいときはAIと人間で差はなさそうだが，書く技能を高めたいときはCEFR_B学習者には教師（人間）のフィードバックが好まれる可能性が高い」ことが分かった。これは松井（2020）の「AIフィードバックに不信感を抱くCEFR_A1～B1学習者はライティングのCEFRレベルアップが見られなかった」という結果に関して，新たな示唆を与える結果となった。



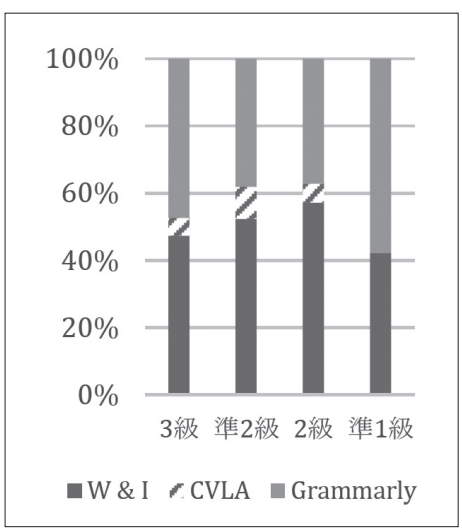
■図 6・図 7「書く意欲」と「書く技能」に関するフィードバックの選択の割合

3.1.3 RQ3：学習者のCEFRレベルで選ぶAIツールに差があるか

表6は事前・事後タスクで取り組んだ英検の級別に，事後アンケート項目6「練習タスクで一番使用したツール」を示したものである（図8はその割合）。CVLAは全級で母数が少なかったので分析対象から除外し，W & IとGrammarlyで比較を行った。直接確率計算を実施した結果，全級で有意差は見られなかった（3級p=0.55，準2級p=0.32，2級p=0.15，準1級p=0.32，片側検定）。RQ3「学習者のCEFRレベルで選ぶAIツールに差があるとは言えない」ことが示唆された。

■表6: 事前・事後タスクの英検級別のAIツールの選択人数 (n=151)

タスク	W & I	CVLA	Grammarly
3級	36	4	36
準2級	11	2	8
2級	20	2	13
準1級	8	0	11



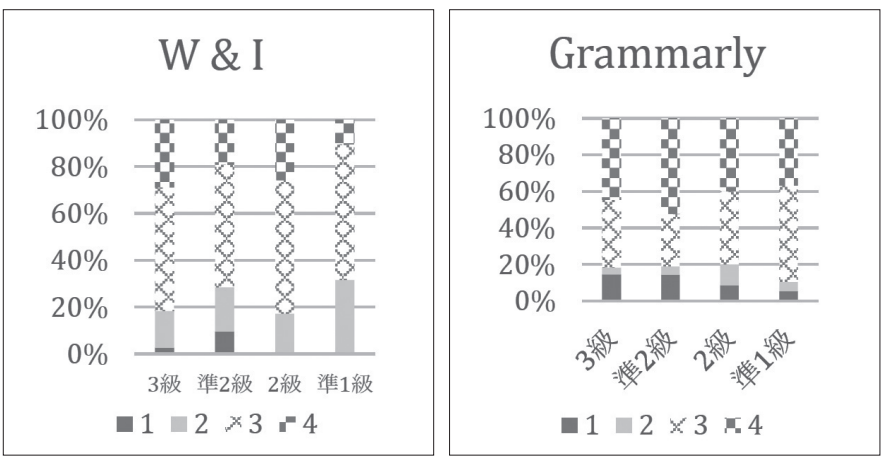
■図8 AIツールの選択の割合

3.1.4 RQ4：学習者のCEFRレベルでAIフィードバックの理解度に差があるか

表7は，事前・事後タスクで取り組んだ英検の級別に事後アンケート項目16, 17「AIフィードバック（W & IとGammarly）の意味（意図）」について【ほどんど分からなかった（1点）】【あまり分からなかった（2点）】【だいたい分かった（3点）】【よく分かった（4点）】の4件法で回答した平均点を示したものである（図9, 図10はその割合）。事前・事後タスクで取り組んだ級別に分散分析を行った結果，有意差は見られなかった（それぞれ$F(3, 147)=1.45$, $p=.23$,）$F(3, 147)=0.07$, $p=.98$）。RQ4「CEFRレベルでAIフィードバックの理解度に差があるとは言えない」ことが示唆された。

■表7: 事前・事後タスクの英検級別のAIフィードバックの理解度の平均点（n=151）

	W & I	Grammarly
3級	3.08	3.12
準2級	2.81	3.19
2級	3.09	3.11
準1級	2.79	3.21



■図9・図10 AIフィードバックの理解度の割合

3.2 RQ5：学習者のCEFRレベルや学年で自律学習能力スコアに差があるか

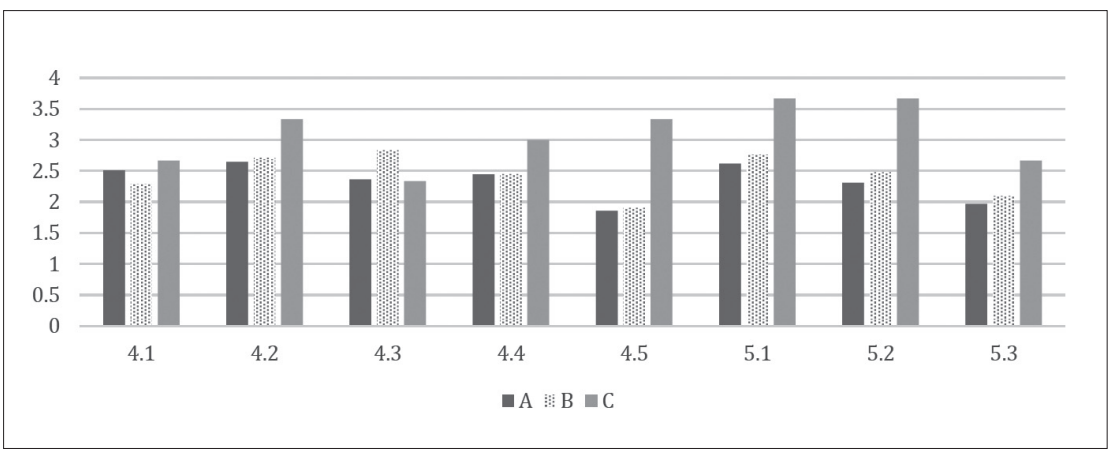
表8, および図11, 図12は遅延タスクに取り組んだ274名の遅延アンケート（CEFRレベルと「自律学習能力」に関する質問8項目（表1の4.1〜5.3））についてCEFRレベルおよび学年別の回答を示したものである。なお，CEFRレベルはA1, A2をCEFR_Aに，B1, B2をCEFR_Bに，C1, C2をCEFR_Cにまとめたが，CEFR_C学習者数は3名と少ないのでCEFR別の分析から除外した。CEFR_A学習者とCEFR_B学習者における項目ごとの平均のt検定の結果，項目4.3「教師の説明がなくても，AIツールのフィードバックを理解することができる」と思う意識がCEFR_A学習者よりCEFR_B学習者で有意に高かった（$p=.00$）。

学年別に見ると，項目4.1「授業時間外に（教師のいないところで）ライティングの学習をするのは不安だ」と思う意識は，1,2年生で5年生より有意に高かった（1年＞5年$p=.002$, 2年＞5年 $p=<.00$）。項目4.3「教師の説明がなくても，AIツールのフィードバックを理解することができる」と思う意識は，1年生より6年生で有意に高かった（1年＜6年 $p=.02$）。項目4.5「ライティングを学習するのに必要な時間を授業時間外でも十分に確保している」と思う意識は，5年生より1, 6年生で有意に高かった（1年＞5年 $p=.03$, 6年＞5年 $p=.04$）。項目5.1「今学習していることを理解しているかどうか自分で確かめることができる」と項目5.2「自

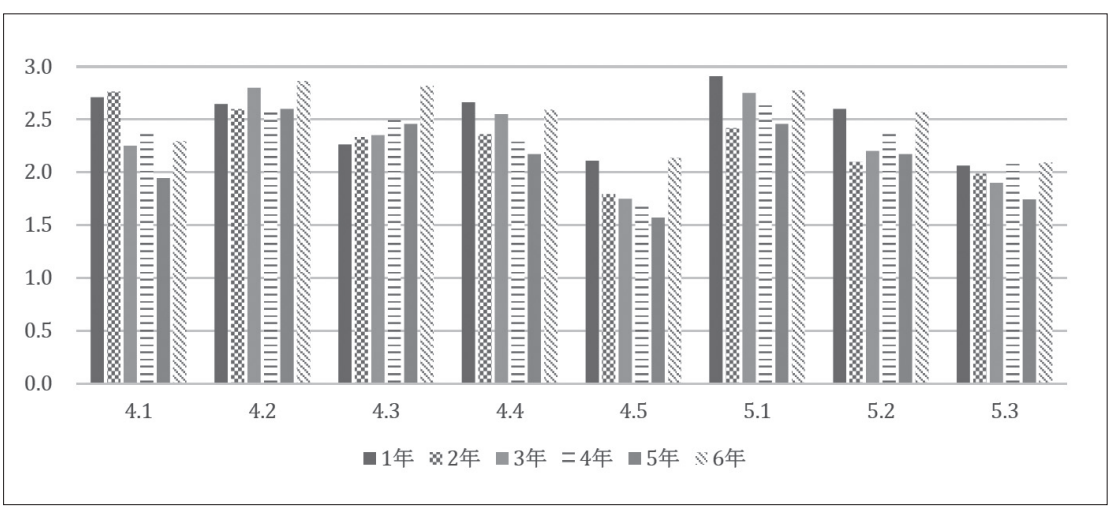
■表8: 遅延タスクのCEFR別および学年別「自律学習能力」に関する質問項目の平均点 (*n*=274)

アカウント	CEFR				学年						
	229	42	3		65	72	20	38	35	44	
	A	B	C		1年	2年	3年	4年	5年	6年	
4.1	2.51	2.29	2.67		2.71	2.76	2.25	2.37	1.94	2.30	**
4.2	2.65	2.71	3.33		2.65	2.60	2.80	2.58	2.60	2.86	
4.3	2.36	2.83	2.33	**	2.26	2.33	2.35	2.50	2.46	2.82	*
4.4	2.45	2.45	3.00		2.66	2.36	2.55	2.32	2.17	2.59	
4.5	1.86	1.91	3.33		2.11	1.79	1.75	1.71	1.57	2.14	**
5.1	2.62	2.76	3.67		2.91	2.42	2.75	2.63	2.46	2.77	**
5.2	2.31	2.48	3.67		2.60	2.10	2.20	2.40	2.17	2.57	*
5.3	1.97	2.10	2.67		2.06	1.99	1.90	2.08	1.74	2.09	

**p=<.01　*p<=.05



■図11 CEFR別「自律学習能力」に関する質問項目の平均



■図12 学年別「自律学習能力」に関する質問項目の平均

分のライティングの実力を自分で評価する方法を知っている」と思う意識は，2年生より1年生で有意に高かった（p=.00, p=.03）。

まとめると，RQ5「CEFRレベル別だと，CEFR_A学習者よりCEFR_B学習者の方が自律学習能力スコアが高く，学習を主体的に展開できる可能性が高い」ことが示唆された。RQ5学年別だと，「中学校低学年では教師のいないところでライティング学習をすることに対する不安意識が強い」ことや「高校3年では教師の説明がなくてもAIフィードバックを理解してCEFRレベルに関係なく個別最適化させた学習を展開できる可能性が高い」ことが示唆された。また，「高校2年では授業時間外でのライティング学習時間の確保の必要性を強く感じる可能性が高い」ことも示唆され，個別学習に最適な学習方法を改めて提示したり吟味したりする機会を設けると良いことが示唆された。「1年生が2年生より自分のライティング力に関して自分で確かめたり自分で評価したりする方法を知っているという意識が高かった」ことは意外な結果であった。

3.3 教員アンケートの分析

教職経験5-10年の教員2名と教職経験20年以上の教員4名の計6名の教員から回答を得た。これは全学年の授業をカバーする。

活動時期について「良い」理由として，「英検受験前にできたから」，一方改善すべき点は「低学年には事前に英検のライティングセクションの指導をすべきだった」「低学年も英検を意識してもう少し早く実施しても良かった」などが挙げられた。

活動回数について「良い」理由として，「授業でライティング活動をあまり行えていなかったから」，一方改善すべき点は，「タブレット操作に不慣れな生徒が回数を重ねても差が開かないようにする」「成果（変化）を感じられる生徒が物足りなさを感じないようにする」「授業時間をそれほど使わないようにする」などが挙げられた。

活動方法について「良い」理由として，「ライティングに集中して取り組むきっかけになったから」「AIの使い方に慣れたから」「ライティング指導は時間がかかるけどAIツールを使うと解決できるから」「解説動画があったおかげで前期生も自分で進めることができたから」などが挙げられた。一方改良すべき点として，「ドキュメントの表記が全て英語だったので（資料4），前期生には日本語表記にする」「前期生にはスモールステップを用意する」などが挙げられた。

主体的学習を展開する効果が「高い」理由として，「ICTツールに慣れている生徒や英語がある程度できる生徒にとっては良い」「授業内外で積極的に使う姿が見られた」「自分が書いた表現で学ぶことができる」「フィードバックがすぐもらえる」などが挙げられた。一方，「間違いの直し方が分からない」「複数のツールを使いこなせない」「主体性を高めるには回数が足りない」などの改善点が挙げられた。

ライティング力を高める効果が「高い」理由として，「語彙力や文法力がある人には有効」「活動を繰り返すことで力がつく」などが挙げられた。一方，「AIは単語・文法・表現の間違いは指摘するが，文章全体の論理性については改善を促せない」「1年生にはレベルが高い」などの意見が見られた。

まとめると，本研究で実施したAIツールを活用したライティングタスクは，中学校低学年にはICTツールやまとまりのあるライティングタスクに慣れさせるために，スモールステップや動画の活用など丁寧な導入が必要だが，一方で，教師が抱えるライティングのフィードバックへの負担を軽減し，AIの瞬時のフィードバックにより学習者の学習機会を豊富に確保できる可能性が示唆された。また，生徒の個別最適な学習を期待するなら，英検や定期考査のタイミングで，AIツールを活用して，自律学習能力を継続的に授業で振り返る活動を取り入れ，論理的展開を指導する段階ではAIツールに頼るだけでなく引き続き人間の教師が介入する必要性が示された。

4 まとめと教育的示唆

本研究は，日本人中高生がAIツールを活用してライティングタスクに取り組んだ時に，主体性が高い学習を展開できるかどうかを明らかにするために実施した。生徒の自力でのCEFRレベルアップを目標に掲げ，フィードバックの特徴が異なる3つの選択肢から最適なAIツールを生徒主体で選び，思考力や判断力を働かせて，教師の支援がほとんどなくてもAIが出した結果（フィードバック）を読み取り，個別学習に活用することをねらって実践した。本研究では「主体性が高い学習」を「自律性や有能性を高めることをねらった学習」と捉え，「自律学習能力」を育成することをねらって学習タスクをデザインした。

RQ1に関して，AIツールを使ってライティングタスクを行っても，どのCEFRレベル学習者もライティングレベルが下がることはなかったが，特にCEFR_A学習者が自力でレベルアップする可能性は低いことが示唆された。この結果は，松井（2020）の「CEFR_A2学習者がW & Iを使ったライティング活動でライティングのレベルが上がった」という結果を支持するものではなかった。また，CEFR_A1学習者は中学校低学年が多く，端末やブラウザ・アプリなどの操作に不慣れな段階のタスク遂行が困難な生徒が半数弱見られたり，教師がいないところでライティング学習をすることへの不安が見られたりした（RQ5）。今後は端末操作等に慣れる生徒が増えると思われるので，低学年に合わせてタスクを細分化・最適化し，手順に慣れさせ，不安を払しょくした上で本タスクを実施したい。

RQ2に関して，AIと人間教師で，どちらが「書く意欲」や「書く技能」をより高めると学習者に思われるかを分析した。松井（2020）は「W & Iを使ったライティング活動でCEFR_A1～B1学習者のライティングレベルに伸長が見られなかった要因はAIへの不信感だ」と結論付けた。本研究では，「書く意欲」に関して，どのCEFRレベル学習者でもAIと人間教師で選択に有意差がなかったことから，書く意欲を高める目的でフィードバックにAIを代用しても支障はないと考える。一方，「書く技能」に関してはCEFR_B1学習者で人間教師のフィードバックの方がAIより高まると回答した人が多かったことや，高校2年生では学習方略への悩みが見取れた（RQ5）。高校2年生のCEFR_B1学習者には，書く技能を高めるためにAIでは補えない人間教師のフィードバックが期待されていることが分かった。CEFR_A学習者には教育AIを活用して書く機会を豊富に確保して書く意欲を高め，CEFR_B学習者には論理の展開など無料のAIツールでは補えない質に関わるフィードバックを人間教師が支援しつつ，主体的かつ個別最適な活動が展開できる可能性が示唆された。

RQ3，RQ4に関して，CEFRレベルの違いで好まれるAIツールやAIフィードバックの理解度に差が出るという結果は出なかった。中学生からAIツールを活用して，最適なAIツールを主体的に選択し，思考力や判断力を働かせながら，AIフィードバックを主体的に読み取ろうと粘り強く取り組ませることが可能なタスクであることが示唆された。

RQ5に関して，本研究では平成29～30年改訂学習指導要領の観点別評価の「主体性」に関して，CEFRの「自律学習能力」に置き換えて分析した。特に，「自らの学習を調整しようとする側面」と「言語活動への取組に関して見通しを立てたり振り返ったりして自らの学習を自覚的に捉えている側面」に注目して，自律学習能力の項目4「学習のコントロール・自己管理」と項目5「評価」を分析した。分析の結果，CEFR_A学習者よりCEFR_B学習者の方が，また1,2年生より6年生で項目4.3「教師の説明がなくても，AIツールのフィードバックを理解することができる」と思う意識が高く，学習を主体的に展開できる可能性が高いことが分かった。一方で，2年生より1年生で項目5.2「自分のライティングの実力を自分で評価する方法を知っている」と思う意識が高く，自律性や有能性を高めるといった主体的な学習を展開することをねらう場合，これは育成すべき期待する姿であるため，中学生から自律学習能力の育成をライティング指導の一環として授業に位置づけたい。5年生では項目4.5「ライティングを学習するのに必要な時間を授業時間外でも十分に確保している」と思う意識が1，6年生より低かったことから，受験勉強という長期的な学習計画を意識する学年でもあるため，改めて個別最適な学習を捉え直す機会を設けたい。

教員へのアンケートからは，AIツールを活用することでライティングのフィードバックへの負担減や豊富にライティングタスクを設定することに期待が示された。また，CEFRレベルごとに生徒のニーズ把握が可視化され，指導デザインのヒントを得たというコメントも見られた。本研究では「自律学習能力」の項目「1. 目標設定」「2. 学習内容の明確化」「3. 学習方法・教材の明確化」は分析対象外であったが，主体性を高めるためには中学校低学年から「自律学習能力」を生徒自身に自覚させ，学習方略を捉え直す学習機会を持たせる必要がある。生徒が自力でCEFRレベルアップに向けて粘り強く学習を行えるよう，徐々に学習の責任を教師から生徒へ委ねるという教育観が必要である。

本研究では，教育AIを利用して，公立学校における「1人1台端末」の有効な活用方法を，外国語教育において探った。分析結果から，教育AIを活用して生徒主体で個別最適化された学習を推進できる可能性が示された。年間のライティングタスクに位置づけ，教科書を使用した活動と並行して実施し，生徒の主体的な学びを推進したい。

謝辞

本研究は公益財団法人 日本英語技能検定協会の研究助成金で実施した。実施に当たって，斉田智里先生から多くの助言を賜った。また，前任校の教師や生徒の皆さんからは熱心な取り組みと振り返りのコメントを頂いた。研究結果は，研究目的以外には使用しないこと，同意なく第三者に提供しないこと等を口頭で説明した。また，研究で扱った生徒等のデータ使用にあたっては，個人が特定されないように統計処理した。研究にご協力くださった方々にお礼申し上げます。なお，本報告書に関して，開示すべき利益相反関連事項はありません。

引用文献

赤堀侃司 (2022).「1人1台端末環境における授業デザインの予備的研究」『AI時代の教育論文誌』第4巻，25-30. 東京都：AI時代の教育学会.

大木充 (2021).「CEFRにおける自律学習の役割とアンリ・オレックの自律学習」，西山教行・大木充（編）『CEFRの理念と現実　現実編　教育現場へのインパクト』，第3章. 33-66 東京都：株式会社くろしお出版.

張建 (2019).「教育AIに関する倫理的課題」『東京電気大学総合文化研究』第17号，31-36. 東京都：東京電機大学.

Koizumi, R., & Nakatsuhara, F. (Eds.) (2022). “Implementing formative and summative classroom assessments of speaking and writing: Promises and challenges of learning-oriented practices.” Language Testing Research Colloquium 2022, Opening Symposium.

国立教育政策研究所 (2021).『「指導と評価の一体化」のための学習評価に関する参考資料（高等学校編）外国語』. 東京都：国立教育政策研究所.

松井市子 (2020).「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ねらって－」EIKEN BULLETIN. 32, 125-143. 東京都：日本英語検定協会.

文部科学省 (2018).『高等学校学習指導要領外国語編英語編』. 東京都：文部科学省.

文部科学省 (2019).『新しい時代の初等中等教育のあり方』. 東京都：文部科学省.

文部科学省(2020)『外国語の指導におけるICTの活用について』. 東京都：文部科学省.

文部科学省(2023)『初等中等教育段階における生成AIの利用に関する暫定的なガイドライン』. 東京都：文部科学省.

森本康彦 (2020).「教育AI・学習記録データを活用した個別最適化された学びのモデル提案」教育システム情報学会第45回全国大会. 講演論文集　第45巻，2-3. 東京都：教育システム情報学会.

向田識弘 (2022).「AIリテラシー教育におけるAIとの向き合い方を考える授業の検討」『AI時代の教育論文誌』第5巻，9-15. 東京都：AI時代の教育学会.

西田理恵子（編）(2022).『動機付け研究に基づく英語指導』. 東京都：株式会社大修館書店.

サラ・マーサー，ゾルタン・ドルニェイ（著）・鈴木章能・和田玲（訳）(2022).『外国語学習者エンゲージメント－主体的学びを引き出す英語授業－』東京都：株式会社アルク.

Zhai, X, et.al. (2021). “A Review of Artificial Intelligence (AI) in Education from 2010 to 2020” Hidawi, Volume 2021, Article ID 8812542, 18 pages, https://doi.org/10.1155/2021/8812542

資料1：グーグルクラスルーム投稿指示

	学習内容
事前タスク	タスクの趣旨説明：AIツールを活用してライティング力アップをねらった活動を行います。全部で6回で，1回目は学習とレベル測定，2～4回目は練習，5回目は効果測定，6回目は自律学習に向けた振り返りです。1回目と5回目に使用するライティングタスクは同じ問題です。AIの力を借りながらライティングレベルアップを図りましょう。 ※W & Iは県支給アカウントでは利用できないようです
	目標：英検の各級のライティング問題と採点基準を知る＋AIツールに慣れる
	教材：①事前テスト問題と解答　②Eiken Writing Format　③W & I　④CVLA　⑤事前アンケート
	1.（5分程度）「ライティングテストの採点に関する観点及び注意点」の解説を聞く 3級：https://www.eiken.or.jp/eiken/exam/2017scoring_3w_info.html ※3級の解説を聴き，他は各自で参照 準2級：https://www.eiken.or.jp/eiken/exam/2017scoring_p2w_info.html 2～1級：https://www.eiken.or.jp/eiken/exam/2016scoring_w_info.html 2.（～20分）取り組みたい級を決め，コードを使ってW & Iのworkbookを開き，解答を作成→レベルチェック＋スクショ W & I：https://writeandimprove.com/workbooks#/wi-workbooks コード：3級 [X9CGZG9Y] 準2級 [RCY8ZBX7] 2級 [TMYF3RCG] 準1級 [K9KMXCBQ] 1級 [9W4MQ48A] 3.（数分）ドキュメント"Eiken Writing Format"を開き，必要事項（氏名等・級・日付）を記入＋英文とスクショを貼付 4.（数分）以下のツールにも英文を貼り付け，レベルチェック＋スクショ CVLA[password:CVLA]：http://dd.kyushu-u.ac.jp/~uchida/cvla.html 5.（10分）ドキュメントに模範解答打ち込み→上記2つのツールにコピペ→結果をスクショ（→進み具体に応じてカット可） ※模範のレベル [W & I / CVLA]：3級 [A2/A1] 準2級 [A2/A2] 2級 [B1/B1] 準1級 [B2/C2] 1級 [C1/C2] 6.（数分）自作と模範の英文末尾にCEFRレベルと語数をドキュメントに追記 7.（数分）「事前アンケート」に回答：https://forms.gle/mMdmF9k111hHJpjH9
練習タスク	目標：AIツールを使いこなしてレベルアップ!かかった時間を計測して回を重ねる毎に短時間で取り組めるようにする ※1コマ目で決めた級に取り組む
	教材：①練習タスク　②Eiken Writing Format　③W & I　④CVLA　⑤Grammarly　⑥振り返りアンケート
	1.（～20分）W & Iのworkbookから決めた級を選び，解答を作成→レベルチェック＋スクショ W & I：https://writeandimprove.com/workbooks#/wi-workbooks ※"Task Timer"→"Start timer"でかかった時間を計測 2.（数分）マイドライブから"Eiken Writing Format"を開き，必要事項（級・日付）を記入して英文とスクショをコピペ 3.（数分）以下のツールに英文をコピペして結果をスクショ＋ドキュメントにコピペ CVLA[password:CVLA]：http://dd.kyushu-u.ac.jp/~uchida/cvla.html Grammarly：https://www.grammarly.com/ 4.（～20分）3つのツール（W & I, CVLA, Grammarly）をフル活用して20分校正 5.（数分）最良英文をドキュメントの(Final Draft)にコピペ→CEFR，語数，時間をFirst DraftとFinal Draftの （W&I:_,CVLA:_,Word No:_,Time:_）に追記 6.（数分）「振り返り」に回答 練習タスク①：https://forms.gle/26dD3pK2FU9vNpqT9 練習タスク②：https://forms.gle/9pogCyWPWvVr4kLVA 練習タスク③：https://forms.gle/bSJ5zpPsEWuf3ktJ7

資料1：グーグルクラスルーム投稿指示

	学習内容
事後タスク	※1コマ目から約2～4週間空ける・模範解答は参照しない・取り組む級を変えない 目標：AIツールの学習効果を測るため自力で解答する!
	教材：①事後テスト問題　②Eiken Writing Format　③W & I　④CVLA　⑤Grammarly　⑥事後アンケート
	1.（～20分）W & Iのworkbookから決めた級を選び，解答を作成→レベルチェック＋スクショ W & I：https://writeandimprove.com/workbooks#/wi-workbooks ※"Task Timer"→"Start timer"でかかった時間を計測 2.（数分）マイドライブから"Eiken Writing Format"を開き，必要事項（級・日付）記入して英文とスクショをコピペ 3.（数分）以下のツールに英文をコピペして結果をスクショ＋ドキュメントにコピペ CVLA[password：CVLA]http://dd.kyushu-u.ac.jp/~uchida/cvla.html Grammarly：https://www.grammarly.com/ 4.（～20分）3つのツール（W & I, CVLA, Grammarly）をフル活用して20分校正 5.（数分）最良英文をドキュメントの（Final Draft）にコピペ →CEFR，語数，時間をFirst DraftとFinal Draftの_（W&I:_,CVLA:_,Word No:_,Time:_）に追記 6.（5分）「事後アンケート」に回答：https://forms.gle/zcsrgepzSxMBuzNY6
遅延タスク	※定期考査のライティングセクション振り返り用（事後タスクから数か月～半年空ける） ねらい：6年間の英語学習を有機的に!生徒自ら目標を定め，その達成に向けて自己調整しながら成長を実感できるように!
	教材：①Can-doリスト　②【前期生】CEFR-Jライティング CAN-DO テスト【後期生】定期考査　③遅延アンケート
	【CEFR-Jライティング CAN-DO テスト】※ロイロノート活用 ※ロイロノート名：定期考査提出箱名：R4_2_End 1. 資料箱「学内共有」→「cefr-j_writing」→「PreA」「A1」「A2」「B1」「B2」の順に，「解答作成→模範解答例を見て赤で〇付け→A評価と思われたら次のレベルにチャレンジ」の手順で取り組む。※AIツールは適宜利用 2. 提出箱に取り組んだレベルまで提出する。 （例：A1.3の途中ならA1の一連のカードを提出する） 3.「遅延アンケート」に回答する：https://forms.gle/wSgEHN9TK5GMDaVX9 【定期考査用】※W & I活用，CEFR_Aレベル（50 words），Bレベル（100 words） 1.（15分）W & Iのworkbook，"R4 School Exam"（コード：N98MNFGX）に定期考査のライティングセクション問題の解答を入力し，15分間校正する。他のAIツールも活用して自力でレベルアップをねらう。 W & I：https://writeandimprove.com/ CVLA[password:CVLA]：http://dd.kyushu-u.ac.jp/~uchida/cvla.html Grammarly：https://www.grammarly.com/ 2.（数分）最良英文を解答用紙に赤ペンで加除したものと事前事後CEFRレベルを記入 （例：事前B1→事後B2）→スクリーンショットしてロイロノート／クラスルームに提出 ※ロイロノート名：定期考査提出箱名：R4_2_End 3.（5分）「遅延アンケート」に回答する：https://forms.gle/wSgEHN9TK5GMDaVX9

資料2：事前・事後タスクで活用した英検ライティングセクション問題

級	実用英語技能検定試験ライティングセクション問題と模範解答	CEFR
3級	【Question】 ・あなたは，外国人の友達から以下のQuestionをされました。 ・Questionについて，あなたの考えとその理由を2つ英文で書きなさい。 ・語数の目安は25〜35語です。 Question: Do you like going to festivals in summer?	[W&I] A2 [CVLA] A1
	【Model answer】 No, I don't. First, it's too hot in summer, so I don't want to go to festivals. Second, festivals are often very crowded. I don't like going to crowded places.	
準2級	【Question】 ・あなたは，外国人の知り合いから以下のQuestionをされました。 ・Questionについて，あなたの意見とその理由を2つ英文で書きなさい。 ・語数の目安は50〜60語です。 Question: Which do you think is better for people, borrowing books from libraries or buying books at stores?	[W&I] A2 [CVLA] A2
	【Model answer】 I think it is a good idea for people to learn how to cool by using the Internet. First, there are many videos on the Internet about cooking. Learning to cook by watching them is easier than by reading cookbooks. Second, they can learn to cook for free. If people take cooking classes, they have to pay for them.	
2級	【Question】 ・以下のTopicについてあなたの意見とその理由を2つ書きなさい。 ・Pointsは理由を書く際の参考となる観点を示したものです。ただし，これら以外の観点から理由を書いてもかまいません。 ・語数の目安は80〜100語です。 Topic: Some people say that it is necessary for people to go to important historical sites in order to understand history better. Do you agree with this opinion? Points: Experience, Motivation, Technology	[W&I] B1 [CVLA] B1
	【Model answer】 I agree that it is necessary for people to go to important historical sites to understand history better. First, people can experience the actual scale of historical sites, such as the pyramids in Egypt. They cannot experience this by seeing these sites online, even with today's advanced technology. Second, visitors can have the opportunity to talk to local people. These people have stories about local history which include important information that is not in books. Such stories can make people want to learn more about history. For these reasons, I think people should visit historical sites.	
準1級	【Question】 ・Write an essay on the given TOPIC. ・Use TWO of the POINTS below to support your answer. ・Suggested length: 120-150 words TOPIC: Should people's salaries be based on their job performance? POINTS: Age, Company profits, Motivation, Skills	[W&I] B2 [CVLA] C2

資料2：事前・事後タスクで活用した英検ライティングセクション問題

級	日本英語技能検定試験ライティングセクション問題と模範解答	CEFR
準1級	【Model answer】 In my opinion, from the perspectives of motivation and company profits, people's salaries should definitely be related to their job performance. To begin with, while standardized salaries for workers in companies today are common, the level of motivation among employees can vary greatly. Rewarding enthusiastic employees who produce better work with higher salaries is not only fair but would also have the wider benefit of motivating other employees. Additionally, the efforts that employees put into performing their work duties well ultimately benefit companies by increasing their profits. One of the responsibilities of a business is said to be the distribution of the profits to those who contribute to its growth. Therefore, to fulfill this responsibility, companies must make sure that salaries match workers' job performance. To conclude, when considering the importance of employee motivation and sharing company profits, I feel that people's salaries should be based on their job performance.	[W&I] B2 [CVLA] C2
1級	【Question】 ・Write an essay on the given TOPIC. ・ Give THREE reasons to support your answer. ・Structure: instruction, main body, conclusion ・Suggested length: 200-240 words Agree or disagree: Genetic engineering will have a positive influence on society in the future.	[W&I] C1 [CVLA] C2
	【Model answer】 Although still in its infancy, genetic engineering will be a tremendous boon to future societies. Specifically, the technology will eliminate genetic diseases, help protect wildlife, and stabilize the global food supply. Genetic engineering has potentially life-changing health applications. For example, the technology has opened the door to the possibility of identifying fatal diseases early or neutralizing genetic defects before birth. In addition to the improvement in individual quality of life, the long-term significance of healthier populations only reinforces the importance of this technology. The environmental implications of genetic engineering are also pertinent. Accelerated by unprecedented industrialization, global warming has become one of many factors behind the dwindling numbers of wildlife species. Genetic engineering, however, could help bring back extinct species and strengthen the genes of endangered ones, which would benefit society greatly as healthy ecosystems are a key source of many resources indispensable for modern needs. Finally, genetic engineering offers modern-day solutions that promise lasting impacts on future generations. Growing populations are putting greater strain on the global food supply, which is already at the mercy of worsening climate change. Genetic modification, however, could boost yields and create more-resilient crops, both of which would prove vital in famine-stricken areas and regions with poor climates. Therefore, tackling human health issues, the issues of endangered species, and escalating food crisis via genetic engineering underscores just how invaluable a tool such technology will be for effecting positive social change in the future.	

資料3: アンケート項目

	アンケート項目
事前タスク	1. 現在の英検取得級【選択肢:1 ～ 3級／その他／なし】 2. 自分は現在英検何級くらいの英語力だと思いますか。【選択肢:1 ～ 3級／その他】 3. 事前タスクで何級に取り組みましたか。【選択肢:1 ～ 3級】 4. その級を選んだ理由は何ですか。【選択肢:自分に適していると思ったから／取り組みたい級だから／なんとなく】 5. 自作英文のCEFR[W & I, CVLA]は何でしたか。【選択肢:A1 ～ C2】 6. 模範英文のCEFR[W & I, CVLA]は何でしたか。【選択肢:A1 ～ C2】 7. W & Iを使った感想を教えてください。【記述式】 8. CVLAを使った感想を教えてください。【記述式】 9. 模範解答と比較して感じたことや分かったことを教えてください。【記述式】
練習タスク	1. 何級に取り組みましたか。【選択肢:1 ～ 3級】 2. 1校目(First Draft)のCEFR[W & I, CVLA]は何でしたか。【選択肢:A1 ～ C2】 3. 最終校(Final Draft)のCEFR[W & I, CVLA]は何でしたか。【選択肢:A1 ～ C2】 4. 一番使ったツールは何ですか。【選択肢:W & I ／ CVLA ／ Grammarly】 5. 4の理由を教えてください。【記述式】
事後タスク	1. 何級に取り組みましたか。【選択肢:1 ～ 3級】 2. 自作ライティングのCEFR[W& I, CVLA]は何でしたか。【選択肢:A1 ～ C2】 3. これまでW & Iを使ってきた感想を教えてください。【記述式】 4. これまでCVLAを使ってきた感想を教えてください。【記述式】 5. これまでGrammarlyを使ってきた感想を教えてください。【記述式】 6. 練習タスクで一番使用したツールは何でしたか。【選択肢:W & I ／ CVLA ／ Grammarly】 7. 6の理由を教えてください。【記述式】 8. 練習タスクで一番使用しなかったツールは何ですか。【選択肢:W & I ／ CVLA ／ Grammarly】 9. 8の理由を教えてください。【記述式】 10. AI添削ツール(W & I, CVLA, Grammarly)を使ってきて感じたことを教えてください。【記述式】 11.「書く意欲」について【選択肢:先生の添削の方が高まる／ AIの添削の方が高まる／どちらも高まる／どちらも高まらない】 12. 11の理由を教えてください。【記述式】 13.「書く技能」について【選択肢:先生の添削の方が高まる／ AIの添削の方が高まる／どちらも高まる／どちらも高まらない】 14. 13の理由を教えてください。【記述式】 15.「主体性」について【選択肢:CEFRをアップするため／他の級・ツールにも授業外でも取り組んだ／取り組まなかった】 16. “W & I” について【選択肢:AIの添削の意味(意図)がよく／だいたい分かった・あまり／ほとんど分からなかった】 17. “Grammarly” について【選択肢:AIの添削の意味(意図)がよく／だいたい分かった・あまり／ほとんど分からなかった】 18. 同じような活動があればまた取り組みたいですか。【選択肢:取り組みたい／取り組みたくない】 19. 18の理由を教えてください。【記述式】

資料3：アンケート項目

	アンケート項目
遅延タスク	属性：遅延ライティングタスク（CEFR-Jライティング CAN-DO テスト／定期考査のライティングセクション）の CEFR[W & I, CVLA] は何でしたか。【選択肢：A1 ～ C2】 ※以下選択肢：そう思う／どちらかと言えばそう思う～どちらかと言えばそう／そう思わない 4.1 授業時間外に（教師のいないところで），ライティングの学習をするのは不安だ。 4.2 教師がいなくても，AIツールがあればライティング力は伸ばせると思う。 4.3 教師の説明がなくても，AIツールのフィードバックを理解することができる。 4.4 私はライティングを学習するための時間をどのように確保したらよいのか分かっている。 4.5 私はライティングを学習するのに必要な時間を授業時間外でも十分に確保している。 5.1 私は今学習していることを理解しているかどうか自分で確かめることができる。 5.2 私は私自身のライティングの実力を自分で評価する方法を知っている。 5.3 私は，時々私自身のライティングの実力を自分で評価している。
教員用	属性：指導歴・担当学年 1. 実施時期について【選択式：良い／悪い／どちらでもない】 2. 1の理由【記述式】 3. 実施回数について【選択式：良い／悪い／どちらでもない】 4. 3の理由【記述式】 5. 実施方法について【選択式：良い／悪い／どちらでもない】 6. 5の理由【記述式】 7. 主体性を高める効果について【選択式：高い／低い／どちらでもない】 8. 7の理由【記述式】 9. ライティング力を高める効果について【選択式：高い／低い／どちらでもない】 10. 9の理由【記述式】 11. 感想等【記述式】

資料4：学習記録データ（ドキュメント）

生徒A【英検ライティングセクション問題準2級選択者】の事前・事後タスク

(Class_No._Name)
(Subject Name)
(Teacher's Name)

(Grade_ Pre Writing) Date_

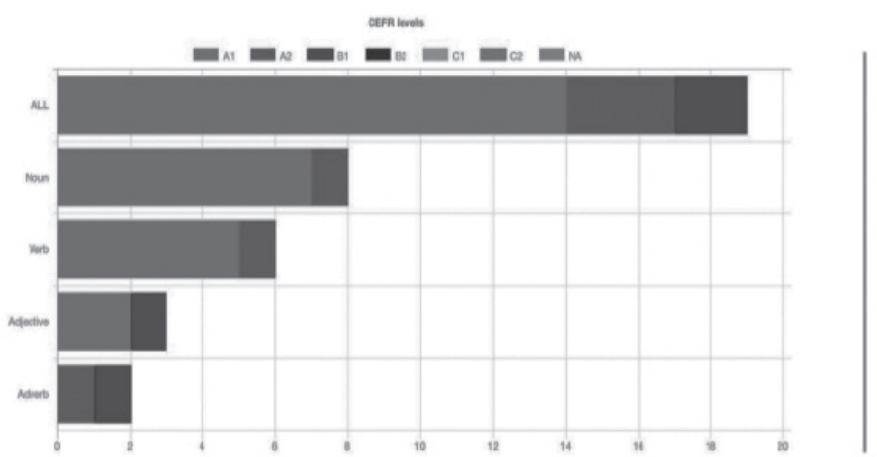
Original Writing

I agree. I have two reasons. The first is that all kinds of information is available on the Internet. You can look it up as you wish. Second, I can look it up on the spot. Therefore, I think it is a good idea to learn cooking on the Internet.

Classroom 9:57 9月14日(水) 69%
安全ではありません — dd.kyushu-u.ac.jp

B1	10.00	2.37	1.57	0.18
B2	12.33	2.88	1.71	0.26
Input	0.16	1.67	1.37	0.12
Estimated level	PreA1	A2.1	A2.1	A2.1

Estimated Text Level:A1.2



(CEFR W & I:A2_, CVLA:A1.2_, Word No:_)

Model Writing

I think it is a good idea for people to learn how to cook by using the Internet. First, there are many videos on the internet about cooking. Learning to cook by watching them is easier than by reading cookbooks. Second, they can learn to cook for free. If people take cooking classes,they have to pay for them. (CEFR W & I:_, CVLA:_, Word No:_)

資料4：学習記録データ（ドキュメント）

生徒A【英検ライティングセクション問題準2級選択者】の事前・事後タスク

(Grade_ Post Writing) Date_

First Draft

I think it is a good idea for people to learn how to cook by using the internet. I have two reasons. First, The Internet can be used to research and learn from many different perspectives. Second, You can learn very clearly by watching the video. Therefore, I think it is a good idea for people to learn how to cook by using the internet.

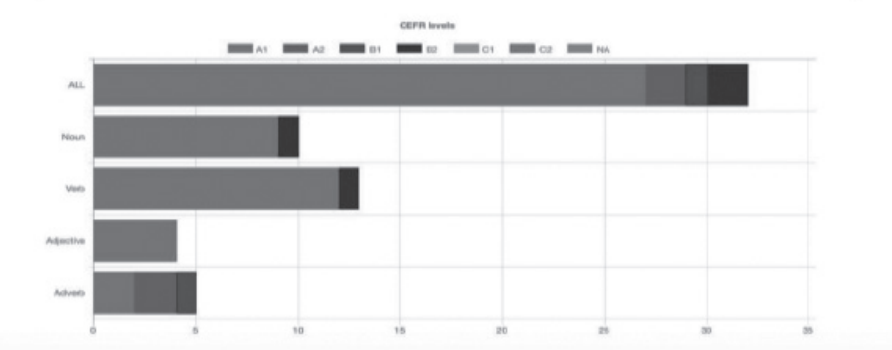
(CEFR W & I:A2 _, CVLA:A2_, Word No:_, Time_)

Final Draft

I think it is a good idea for people to learn how to cook by using the internet. I have two reasons. First, The Internet can be used to research and learn from many different perspectives. Second, You can learn very clearly by watching the video. Therefore, I think it is a good idea for people to learn how to cook by using the internet.

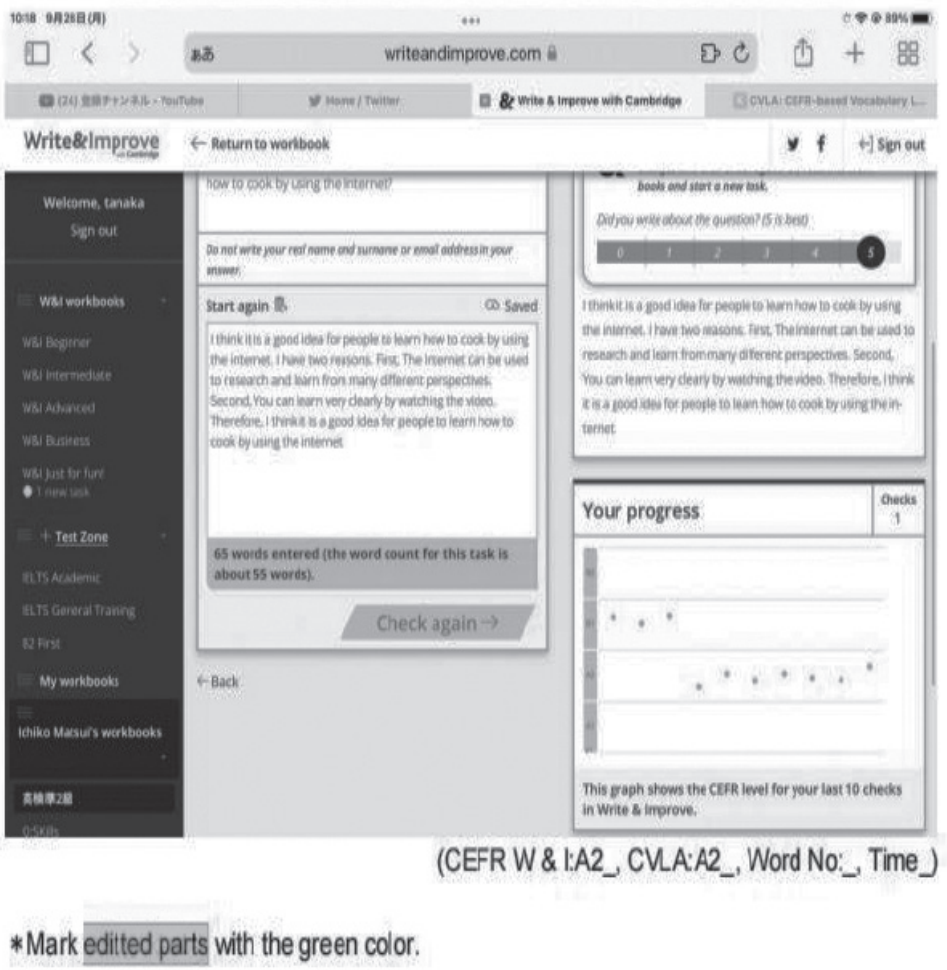
B1	10.00	2.37	1.57	0.18
B2	12.33	2.88	1.71	0.26
Input	4.34	3.40	1.31	0.10
Estimated level	A1.1	C1	A1.2	A2.1

Estimated Text Level:A2.2



資料4：学習記録データ（ドキュメント）

生徒A【英検ライティングセクション問題準2級選択者】の事前・事後タスク



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

学習者の英語観と情意要因

― グローバル英語に焦点を当てて ―

研究者:池田　幸恵　広島県／広島商船高等専門学校　准教授
（申請時:広島県／広島商船高等専門学校　講師）
《研究助言者:斉田 智里》

概要　本研究は，高専生を対象として，母語話者英語とグローバル英語に対する意識の違い，さらに英語観が動機づけを含めた情意要因とどのように関連するかについて検討することを目的とした。質問紙調査の結果，英語観には「母語話者英語志向」，「アイデンティティ」，「グローバル英語志向」の3因子が見出され，母語話者英語志向がアイデンティティおよびグローバル英語志向よりも有意に高いことが示された。一方で，母語話者英語志向とグローバル英語志向の間には中程度の相関関係が示され，母語話者英語志向とグローバル英語志向を対立する概念とは捉えていないという高専生の英語観の特性が示唆された。さらに，英語観が動機づけ，英語完璧主義，スピーキング不安に及ぼす影響について検討した。その結果，母語話者英語志向，アイデンティティ，グローバル英語志向のいずれもが動機づけに対して正の影響を与えているが，グローバル英語志向が最も強く影響していることが示された。またアイデンティティはスピーキング回避に負の影響を及ぼす一方で，英語完璧主義には正の影響を与え，さらに英語完璧主義は動機づけに正の影響を与えていることが示された。これらのことから，グローバル英語志向と日本人英語話者としてのアイデンティティを涵養していくことが重要であることが示唆された。

1 はじめに

1.1 問題と目的

近年，英語使用の拡大に伴って，英語はイギリス，アメリカ，カナダ，オーストラリアなどの特定の国の人々の母語として，母語話者間でのコミュニケーションのために使用される言語という役割を超え，英語を母語としない者同士の間での共通言語として使用されている。このように英語を世界の共通言語とする英語の役割の拡大は，英語についての新たな概念を生み出した。国際共通英語論の先駆者であるKachru（1985）は，英語を第一言語かつ母語とするアメリカなどの内心円圏（Inner Circle）以外の，英語を公用語や準公用語として使用するシンガポールなどの外円圏（Outer Circle）の多様な英語変種に対しても英語としての正当性を付与しようとする「世界諸英語」（World Englishes）という概念を提唱した。また，Smith（1983）によって，非母語話者同士での英語使用という目的の多様性を重視し，非母語話者間のコミュニケーションの道具である英語の多様な英語変種を認めようとする「国際語としての英語」（English as an International Language）という概念が打ち出された。さらに，Jenkins（2000）は，多様な英語変種を認めながらも，非母語話者同士が理解可能になるための基盤となる英語の共通性に着目した「国際共通

語としての英語」(English as an Lingua Franca)という概念を提唱した。こうした新たな英語概念は総称として「グローバル英語」(Global Englishes)と呼ばれている。「英語」を単数形のEnglishではなく複数形のEnglishesで表現することによって，英語をただ一つの標準によって成り立つ特定の話者が所有権を持つ言語としてではなく，様々な非母語話者によって使用されるという話者の多様性と，母語話者英語以外の英語変種を誤りとしない言語の多様性を重視する考え方を打ち出している。

このような英語概念の変化は英語学習者にどのような影響を及ぼすのだろうか。これまでの研究では，日本人EFL学習者の多くが英米の母語話者英語を好み，高く評価する一方で，非母語話者英語を低く評価し，否定的に捉えていることが報告されている(Matsuura et al., 1994; Chiba et al., 1995; Matsuda, 2003; 花元, 2010; Saito, 2014; 松浦・若生, 2019)。そうだとするならば，英語の概念が英米の母語話者英語から非母語話者の多様な英語変種へと変化したことによって，学習者の英語観には変化が生じたのだろうか。そして，学習者がどのような英語観を持つかということは，どのような英語を習得することを目指すかという目標や，目標とする英語に照らし合わせて相対的に自分の英語をどのように捉えるかという自己肯定感や劣等感、英語を習得することにどのような動機や必要性を見出すかといった英語学習の情意要因に関連するのではないかと考えられる。しかしながら，日本人英語学習者の母語話者英語に対する意識や言語態度については知見が蓄えられている一方で，英語概念の変化に伴う学習者の英語観が情意要因にどのような影響を与えるかはこれまで充分に検討されていない。世界的な英語使用の拡大は英語を教室でEFLとして学ぶ日本人学習者にとっても無縁ではなく，検討が必要な課題である。

1.2 日本の学校英語教育における英語観に関する先行研究

日本の学校英語教育では，教科書、教師および学習者それぞれにおいて英語母語話者志向の英語観があることが報告されている。Kawashima(2018)では，1999年，2006年，2016年と縦断的に高校の英語検定教科書指導用CD吹込み者を調べた結果，2016年においても，アメリカおよびカナダの北アメリカ出身者が8割以上であるのに対し，非母語話者は3.8%しかいなかったことを明らかにした。さらに川又(2007)は，同様に高校の英語検定済教科書と教科書準拠CDの吹込み者を調査したところ，オーストラリアやニュージーランド出身という設定の登場人物であるにもかかわらずアメリカ人によって録音されていたケースが多くあり，「見せかけの多様性」であると指摘している。

英米の英語志向は高校教科書だけでなく，音声指導が中心となる小学校英語の教科書にも見られる。教科書に登場する人物には多様性がみられるが，その登場人物がアメリカ人でない場合でも音声はアメリカ英語で拭き込まれており，「音声を多様化することはできていない。というよりもむしろ，英語の音声を多様化しようとはしていない」ことが指摘されている(大石, 2022, p.114)。このような英語学習初期から中期の段階にわたって，アメリカ英語志向による音声教材が教室で使用されていることは、目指すべきはアメリカ英語であり，世界の様々な人々はアメリカ発音の英語を話しており，アメリカ英語こそが標準英語として唯一の正しい英語であるという考え方を学習者に内在化させ強化してしまう可能性があるだろう。

英語教科書だけでなく日本人英語教師にもまた，母語話者英語志向の英語観があることが明らかにされている(e.g., Mimatsu, 2011)。中井(2009)では，中学・高校の日本人英語教師を対象として，発音指導や音読指導のとき，どのような英語変種を使用しているかについてアンケートを実施した結果，アメリカ英語が68.9%，イギリス英語が2.9%，意図的なアメリカ英語とイギリス英語の使い分けが6.2%，英語変種を意識していないが18.7%、その他が3.3%で、日本人英語教師の多数がアメリカ英語を音声指導の学習モデルとしている実態を明らかにした。また，行森(2015)では高校の英語科教員がどのような英語観を有しているかを明らかにしている。英語でのコミュニケーション実践において，英語母語話者と非英語母語話者への意識には違いがあるのかを検討するための調査を実施したところ，英語母語話者への意識が非英語母語話者への意識よりも有意に高い結果となり，教師自身が英語でコミュニケーションを取るとき

の相手，生徒に教える際にどのような相手とコミュニケーションを図れる英語力を目指すかという目標，教師が持つ興味や関心のいずれにおいても非英語母語話者よりも英語母語話者を重視していることを報告している。さらに，英語教師が有する母語話者志向の英語観の要因について因子分析を行った結果、母語話者英語を規範とする「規範主義」、英語圏への「文化的関心」，英語特有の表現や語彙への興味という「言語学的関心」，英語話者との「コミュニケーションへの関心と成功体験」，コミュニケーション力の育成を重視した「実用志向」，知識を教授することを重視した「知識志向」という6つの因子を見出している。

このように日本の学校英語教育では，教科書においても英語教師においても英語母語話者志向が根付いており，そうした環境で学ぶ日本人英語学習者にも英語母語話者志向が見られることが多くの研究によって報告されている。そしてこれらの研究では，英米の母語話者英語は好意的に評価されている一方で，非母語話者英語は否定的に評価される傾向があること，日本人のアクセントによる英語は低く評価されることが指摘されている（Matsuura et al., 1994; Chiba et al., 1995; Matsuda, 2003; 花元，2010; Saito, 2014）。例えば，Tokumoto and Shibata（2011）では，日本，韓国，マレーシアの大学生を対象として，自らの英語アクセントに関するアンケートを実施した。その結果，マレーシア人大学生が最も自身のアクセントを高く評価しているのに対して，日本人グループは自らのアクセントを最も低く評価していた。そして，こうした国による違いについて，その国の歴史的経緯や社会的背景の観点から考察し，教室で外国語として英語を学習する日本では母語話者英語を規範とすることによって，自らのアクセントを劣った直すべきものであるという考えが現れていることを示唆した。

日本人英語学習者を対象とした言語態度の研究はほとんどが大学生を対象としているが，数少ない高校生と中学生を対象とした研究においても母語話者英語を意識した態度が指摘されている。Matsuda（2003）は，高校生を対象とした調査の結果，多くの学生が英語は国際語であることを認識はしているが，英語は英米の母語話者言語であると捉え，そのような英語の習得を目標としている一方で，非英語母語話者には関心を示さず，日本人の英語を正しくないものとして否定的に捉えていることを報告している。中学生を調査対象とした松浦・若生（2019）でも，中学生と大学生の比較分析の結果，どちらの対象者においてもアメリカ英語に対しては好意的な高い評価である一方で，シンガポール英語に対しては否定的な低い評価となる結果が示された。アメリカ英語志向は中学校という学習初期の段階ですでに形成されており，高校を通じてその傾向が強化され，大学でもその傾向が継続する可能性が指摘されている。

以上のように、これまで中学校，高校，大学の学生を対象とした研究が行われ知見が蓄えられてきた。しかしながら高等専門学校（以下、高専）の学生を対象とした研究はほぼ見られない。高専生は年齢としては高校1年生から大学2年生に相当するが，実践的技術者を目指し早期から理工系専門科目を学ぶ高専生を高校生や大学生と同等として扱うことはできないだろう。まず、高専生は大学受験をしないため，受験を意識した英語知識の重視や規範的英語を必要としない。次に，実践的技術者を育成することを目的とする高専ではアジア圏を中心とした非英語圏の人々との国際交流が主となっている。例えば、高専への留学者はタイ，モンゴル，マレーシアなどのアジア圏からがほとんどで，欧米からの留学生はほぼいない。高専が学術・技術交流の提携をしている学校も多くをアジア圏が占めており，高専生が海外研修に行く場合もアジア圏であることが多いため，高専生は英語をアメリカ人やイギリス人よりもアジア圏の人々とのコミュニケーションのために使用する機会が豊富である。これらのことから，高専生はグローバル英語に親和性がある学習環境にいると言うことができる。その高専生がどのような英語観を有しているのかを検討することによって，これまで蓄えられてきた日本人英語学習者の英語観に関する知見に高専生という新たな対象者の知見を加えることには意義があると考えられる。

1.3 英語観と情意要因に関する先行研究

これまで主に大学生を対象者として，日本人英語学習者がどのような言語態度や英語観を有しているかを明らかにしようという研究が取り組まれてきた。しかしながら，そのような英語に対する態度や意識，

特に新たな英語概念であるグローバル英語としての英語観を持つことが英語学習の動機づけを含めた情意要因とどのように関連しているのかという実証研究は十分には行われてきていない。数少ない研究の一つであるMunezane（2013）では，グローバル英語に対する価値観は「英語学習者の動機づけの心理的な景色の中で重要な役割を果たす」と述べ，動機づけとの関連について検討を行った（p. 153）。日本人大学生を対象として，動機づけモデルを共分散構造分析によって分析した結果、グローバル英語に対する価値観が動機づけの予測要因であることを明らかにした。また、Takahashi（2021）は，グローバル英語に対する意識として「英語の広がりに対する肯定的感情」，「英語の広がりの実用的側面」，「異文化コミュニケーションの道具としてのグローバル英語」という3つの因子を見出し、それらの因子と動機づけの相関を分析したところ，動機づけと「実用的側面」の間にr = .44と最も強い相関があり，次いで動機づけと「肯定的感情」の間にr = .36、動機づけと「コミュニケーションの道具」の間にr = .27で正の相関関係があったことを報告している。そしてこの結果から，学習者は英語学習がもたらす経済的な結果を意識するときに，英語を学習しようと動機づけが高まる可能性を示唆した。

これらの研究ではグローバル英語に対する意識と動機づけの関連性についてのみ検討されている。しかし，英語を母語話者英語としてではなくグローバル英語として意識することは，他の情意要因とも関連するのではないかと考えられる。母語話者英語は唯一の標準英語を目標として，正確さを重視するのに対して，グローバル英語は日本人英語を含め多様な英語変種を容認し、正確さよりも理解できること（intelligibility）を重視する英語概念であることから，自分の英語能力に対する不安や自分の英語が他者から評価されることへの不安などの「外国語学習不安」や，英語の習得に完璧さを求める「英語完璧主義」といった英語学習の情意的側面とも関連するのではないかと考えられる。

不安と外国語学習の間には負の関係があることが明らかにされている（e.g., Oxford, 1999）。学習者は外国語学習に不安を感じると，動機づけや学業成績が低下することが報告されている（Yashima et al., 2008; Dörnyei & Ushioda, 2011）。学習者は外国語学習における不安をリスニングとスピーキング活動に対して，特にスピーキングに対して最も強く感じていることが指摘されている（Horwitz, 1986）。Price（1991）は，クラスメートの前で外国語を話すことに学習者が強い不安を抱くのは、発音を間違えること，間違いをクラスメートから馬鹿にされること，コミュニケーションが上手く取れるかを心配しているためであると述べている。スピーキング活動におけるこうした不安は，学習者が母語話者英語志向の英語観を有していることが一因となっている可能性が考えられる。正しい英語を使用しなくてはならないという意識が間違いを過度に恐れることにつながり，英米の英語変種だけを標準英語と捉えることで、標準英語に到達できない自身の英語に劣等感を持ってしまいコミュニケーションが上手くいくか不安になり，日本人英語を話す自己の英語を他者から評価されることに不安を抱くというように，母語話者英語志向の英語観とスピーキング不安の間には関連性がある可能性が推察される。この点について明らかにするためには，学習者が有する英語観が母語話者英語志向かグローバル英語志向かという違いによって，学習者が持つスピーキング不安には違いがあるのかを検討する必要があるだろう。

さらに，母語話者英語志向の英語観は英語学習における完璧主義とも関連すると考えられる。母語話者英語を志向することは母語話者英語を唯一の正しい標準英語として捉え，それ以外の学習者の中間言語（Interlanguage）を含め，母語話者英語以外の英語変種を不完全な望ましくないものと見なすことに結びつくと考えられる。そのため，母語話者英語志向の英語観は英語学習において，間違いのない正確な英語を話そうとすることや，英語母語話者のように完璧に流暢な英語を使用できることを目指そうという英語完璧主義を希求するのではないかと推察される。

また，外国語学習不安は完璧主義と関連性があることが指摘されている。Gregersen and Horwitz（2002）は初めてこの関連性を明らかにした研究であり，Horwitz（1986）による外国語学習教室不安尺度（FLCAS, the Foreign Language Classroom Anxiety Scale）を用いて，不安得点が高い学生と低い学生にグループ分けし，事前に撮影した自分たちのプレゼンテーションの録画を見ているときの様子を完璧主義の指標を用いて分析した。その結果，不安が高い学生と低い学生では全く異なる反応をし，不安が高い学生は

完璧主義の傾向を示していたことを報告している。さらに、Dewaele（2017）では，日本人大学生を対象に外国語学習教室不安と完璧主義についての質問紙調査を実施した結果，外国語学習教室不安と完璧主義の下位尺度である「間違いへの懸念」との間（r = .23）および「自分の行動についての疑い」との間（r = .29）に有意な相関があったことが示されている。母語話者英語志向，グローバル英語志向，英語完璧主義，そしてスピーキング不安がどのような影響関係にあるかを実証することはこれまで未検討な課題であり，英語学習の情意要因についてより良く把握するために意義があると思われる。

1.4 本研究の目的

本研究では，これまで検討されてこなかった高専生を対象として，高専生が母語話者英語やグローバル英語に対しどのような意識を持っているのかを調査し，高専生の英語観について検討する。そして，英語観が動機づけやスピーキング不安，英語完璧主義とどのように関連しているのかを分析することで，英語観が情意要因に及ぼす影響について検討することを目的とする。本研究がこれらのことを明らかにすることによって，日本人EFL学習者にとって英語の概念をどのように捉えることが英語学習において望ましいのかを考えるための一助となることが期待される。そこで本研究では、以下の2つのリサーチ・クエスチョンを設定した。

RQ1 高専生の母語話者英語とグローバル英語に対する意識には違いがあるのか。

RQ2 英語観は動機づけ，スピーキング不安，英語完璧主義といった情意要因にどのような影響を及ぼすのか。

2 調査方法

2.1 調査対象者

国立高等専門学校の本科1～3年生合計347名（1年生123名，2年生123名，3年生101名）を対象とした。高専は15歳から5年間一貫での技術者育成を教育目標とする高等教育機関である。高専生の多くは，卒業後，専門性を活かした職業に就く。英語に対しては苦手意識が強く，英語力は低迷している。大学1年生に相当する高専4年生のTOEIC® Listening & ReadingのIPテストの全国平均点を表1に示す。文部科学省の学校基本調査によると，国立，公立，私立を合わせた全高専の入学者数は毎年10,000名程であることから，全国の本科4年生の約7～8割に相当する受験者数の平均点となっている。平均点は350～360点台の横ばいが続いている。このことから，本調査の対象者の英語力についても同程度であると思われ，充分に高いレベルにあるとは言い難い状態であると推測される。

■表1: 全国の高専4年生のTOTIC® IPテストの平均点

	2016年	2017年	2018年	2019年	2020年	2021年	2022年
平均点	349	351	351	360	351	364	375
受験者数	8,718	8,364	8,042	7,724	7,307	7,312	8,008

（一般社団法人国際ビジネスコミュニケーション協会『TOEIC® Program DATA & ANALYSIS』2016年版から2022年版までを基に作成）

2.2 調査項目

本研究では，質問紙調査によって，英語観，スピーキング不安，英語完璧主義，動機づけについての調査を実施した。以下に質問紙の内容を示す。

(1) 英語観

Nakayama (2015) による尺度を使用した。この尺度は日本人EFL学習者の学習環境に適用可能な質問項目となるよう作成されており，本調査の対象者にも利用可能だと判断した。この尺度は以下の3つの下位尺度から構成される。1つ目は，「君の英語の発音はアメリカ人みたいだねと英語を母語とする人からほめられたらうれしい」などの項目を含む，英語母語話者の発音に対する意識から成る英語母語話者志向についての「ネイティブ・スピーカー神話」。2つ目は，「日本の学校では，日本人らしい英語の発音を教えるべきだ」といった項目を含む，学習者自身の発音に対する認識から成る「アイデンティティ」。3つ目が，「授業では，アメリカだけでなく様々な国の人の発音による英語教材を使うべきだ」などの項目を含む，様々な国の発音に対する意識から成るグローバル英語志向についての「国際語としての英語」。「アイデンティティ」と「国際語としての英語」は質問項目数が少なく，学習者の意識を捉えきれない懸念があったため，Nakamura et al. (2018) による国際共通語としての英語に関する認識尺度から，日本人学習者に適していると思われる質問項目を「アイデンティティ」に1項目（「自分の英語が日本人なまりの発音だと言われても気にならない」），「国際語としての英語」に1項目（「インド英語やシンガポール英語など様々な国の人の英語の発音を理解できるべきだ」）を付け加えた。「ネイティブ・スピーカー神話」が4項目，「アイデンティティ」が4項目，「国際語としての英語」が3項目の計11項目から構成された。質問項目は，1年生でも理解しやすいよう原典の「ネイティブ・スピーカー」という文言を「アメリカ人のような英語を母語とする人」に，「様々な国の人」という文言を「アメリカだけでなく様々な国の人」へ修正した。1:「全くそう思わない」～6:「とてもそう思う」の6件法で回答を求めた。（資料1参照）

(2) スピーキング不安

石原 (2020) による、磯田 (2008) の尺度を参考にして作成されたスピーキング不安に関する尺度を使用した。「他者からの評価に対する不安」（例:「私は，クラスの人に英語の発音が下手だなと思われないか心配です」），「否定的な能力認知」（例:「私は，自分が考えていることを英語で相手に伝えることができないと思います」），「英語を話すことの回避」（例:「私は，人と英語で話すことは避けたいです」）の3因子を想定した計16項目から構成される。石原 (2020) と同様に，1:「全くあてはあらない」～6:「とてもよくあてはまる」の6件法で回答を求めた。（資料2参照）

(3) 英語完璧主義

関谷 (2009) の英語学習観についての尺度で用いられていた「英語完璧主義」を使用した。「英語完璧主義」は，「英語ができるようになるということは，どんな英語でも正確に日本語に訳せるようになることだ」などの7項目から構成される。関谷 (2009) では5件法であったが，他の尺度が6件法であることから統一した方が回答者の混乱を招かず負担を減らせると考え，1:「全くあてはまらない」～6:「とてもよくあてはまる」の6件法で回答を求めた。（資料3参照）

(4) 動機づけ

Taguchi et al. (2009) の「動機づけ」を日本語に翻訳したものを使用した。「私は英語学習を一生懸命にやっている」などの4項目から構成される。Taguchi et al. (2009) と同様に，1:「全くあてはあらない」～6:「とてもよくあてはまる」の6件法で回答を求めた。（資料4参照）

2.3 実施手順

本調査は2023年5月に実施した。各クラスの英語の授業において，調査者が調査の趣旨について説明した。アンケートは研究のために用いられ，それ以外の用途で用いられないこと，データはすべて統計的に処理・分析され，個人が特定されることはないこと，アンケートの回答結果は授業の成績とは一切関係しないことを説明した上で，本調査への参加に同意した学生にアンケートへの回答を依頼した。

質問紙への回答はMicrosoft 365のアンケート・ツールMicrosoft Formsを使用し，学生は各自が所有するスマートフォンから回答した。

本調査から得られたデータは，IBM SPSS Statistics Version 28およびIBM SPSS Amos Version 28を使用して分析を行った。

3 結果

3.1 RQ1についての分析結果

3.1.1 英語観尺度の因子構造の検討

英語観についての質問紙調査から得られたデータについて探索的因子分析（主因子法・プロマックス回転）を行った。すべての項目で天井効果およびフロア効果は見られなかったため，すべての項目を使用した。固有値1以上を基準として，因子の解釈可能性を考慮した結果，3因子解が妥当であると判断した。因子分析の結果を表2に示す。Nakayama（2015）と同様の因子構造が再現された。因子の解釈と命名は先行研究に倣いつつ，本研究の目的に合わせて命名の一部を適宜変更した。第1因子は英語母語話者の発音に関する項目が高い因子負荷を示したことから「母語話者英語志向」と名付けた。第2因子は日本人英語話者としての自身の発音についての項目が中心となっているため「アイデンティティ」と命名した。第3因子は多様な国々の英語の発音を容認する項目が高い負荷量を示していたことから「グローバル英語志向」と名付けた。クロンバックα係数は，母語話者英語志向でα = .82，アイデンティティでα = .76，グローバル英語志向でα = .83であり，一定の信頼性があると判断した。

■表2: 英語観の探索的因子分析の結果

	項目内容	Ⅰ	Ⅱ	Ⅲ	共通性
F1：母語話者英語志向（α=.82）					
9	私はアメリカ人のような英語を母語とする人の発音を身につけたい。	.86	−.03	−.04	.70
4	日本人はできるだけアメリカ人のような英語を母語とする人の英語の発音に近づくように努力するべきだ。	.77	−.05	.03	.62
5	「君の英語の発音はアメリカ人みたいだね」と英語を母語とする人からほめられたらうれしい。	.73	.01	−.02	.52
1	授業では，アメリカ人のような英語を母語とする人の発音による英語教材を使うべきだ。	.52	−.05	.12	.35
F2：アイデンティティ（α=.76）					
2	日本の学校では，日本人らしい英語の発音を教えるべきだ。	−.14	.73	−.01	.53
8	日本人らしい英語の発音ができる人を尊敬すべきだ。	.04	.71	.15	.59

	項目内容	I	II	III	共通性
10	自分の英語が日本人なまりの発音だと言われても気にならない。	−.07	.64	−.08	.39
6	リスニングでは，英語力のある日本人の発音による英語教材を使うべきだ。	.26	.60	−.07	.41
F3：グローバル英語志向（α=.83）					
7	授業では、アメリカだけでなく様々な国の人の発音による英語教材を使うべきだ。	.01	.01	.90	.81
3	リスニングでは、アメリカだけでなく様々な国の人の発音による英語教材を使うべきだ。	−.02	−.02	.90	.76
11	インド英語やシンガポール英語など様々な国の人の英語の発音を理解できるべきだ。	.18	.18	.47	.42
	因子相関 I	−	.06	.58	
	II		−	.27	

3.1.2 母語話者英語志向とグローバル英語志向の差の検討

高専生の母語話者英語とグローバル英語に対する意識の差異について検討を行うために，各因子の平均値を算出して下位尺度得点として，一要因分散分析（対応あり）を行った。平均値，標準偏差，一要因分散分析の結果を表3に示す。英語観要因には有意差が見られた（F[1.8, 630] = 40.27, p < .01）。Bonferroniの方法による多重比較の結果，英語観の3つの因子間すべてにおいて有意差があり，母語話者英語志向はアイデンティティおよびグローバル英語志向よりも有意に高いことが示された。

■表3: 英語観の平均値，標準偏差，分散分析結果

英語観	*M*	*SD*	*F*	多重比較
母語話者英語志向	4.58	1.10	F(1.8, 630) = 40.27**	母>グ>ア
アイデンティティ	3.93	1.04		
グローバル英語志向	4.21	1.22		

**p<.01　母：母語話者英語志向、グ：グローバル英語志向、ア：アイデンティティ

3.2 RQ2についての分析結果

3.2.1 スピーキング不安尺度の因子構造の検討

■表4: スピーキング不安の探索的因子分析の結果

	項目内容	I	II	III	共通性
F1：他者からの評価不安（α=.90）					
7	私は，クラスの人に英語を話すのが下手だなと思われないか心配です	.91	−.03	−.06	.74
1	私は，クラスの人に英語の発音が下手だなと思われないか心配です	.84	−.04	−.13	.58
13	私は，私の話す英語がクラスの人に上手く伝わっているかどうか心配です	.77	.16	−.10	.66
2	私は，先生に英語のスピーキング力を評価されているとき，不安になります	.73	−.11	.17	.59

	項目内容	Ⅰ	Ⅱ	Ⅲ	共通性
4	私は，クラスの人の前で英語を話すとき，緊張します。	.71	.00	.01	.52
5	私は，私の話す英語が先生に上手く伝わっているかどうか心配です	.66	−.05	.18	.55
10	私は，先生が私の英語の間違いをチェックしているときに不安になります	.66	.07	.07	.55
15	私は，クラスの人が英語を上手に話しているときに不安になります	.65	.06	.00	.47
F2：自己の能力不安（α=.88）					
8	私は，英語を話す力がないと思います	−.04	.82	.09	.72
11	私は，今の英語力では英語で話すことはできないと思います	−.01	.81	.01	.66
14	私は，自分が考えていることを英語で相手に伝えることができないと思います	−.06	.79	.07	.65
16	私が話す英語は、相手に意味が伝わらないと思います	.16	.73	−.07	.61
F3：スピーキング回避（α=.83）					
3	私は，できれば人と英語で話したくありません	−.03	−.07	1.02	.92
6	私は，人と英語で話すことは避けたいです	.08	.05	.77	.52
12	私は，英語で話さなければならないときはできるだけしゃべらないようにしています	.05	.30	.46	.65
9	私は，英語で話さなければならないときでも，日本語で話したいと思うことがよくあります	−.07	.19	.41	.26
	因子相関　Ⅰ	−	.56	.53	
	Ⅱ		−	.64	

スピーキング不安についての質問紙調査から得られたデータについて探索的因子分析（主因子法・プロマックス回転）を行った。すべての項目について天井効果とフロア効果の有無を確認したが，いずれも見られなかったため，すべての項目を使用した。固有値1以上を基準として，因子の解釈可能性を考慮した結果，3因子解が妥当であると判断した。因子分析の結果を表4に示す。先行研究のスピーキング不安の構成要素とおおむね整合している因子構造が得られた（注1）。第1因子は他者からどのように評価されるかという項目に高い因子負荷量を示したため「他者からの評価不安（評価不安）」と名付けた。第2因子は自分のスピーキング能力に対する疑いの項目が中心となっているため「自己の能力不安（能力不安）」とした。第3因子は英語で話すことを避けようとする項目に高い因子負荷量を示していたことから「スピーキング回避（回避）」と命名した。クロンバックα係数は、評価不安で $\alpha = .90$，能力不安で $\alpha = .88$，回避で $\alpha = .83$であり，一定の信頼性があると判断した。

3.2.2 各尺度の記述統計量

英語観，スピーキング不安，英語完璧主義，動機づけの記述統計量とクロンバックα係数の結果を表5に示す。英語完璧主義は $\alpha = .87$，動機づけは $\alpha = .86$だったため，一定の信頼性があると判断した。

■表5: 各尺度の記述統計量とα係数

	M	*SD*	α係数
母語話者英語志向	4.58	1.10	.82
アイデンティティ	3.93	1.04	.76
グローバル英語	4.21	1.22	.83
評価不安	4.26	1.04	.90
能力不安	3.69	1.11	.88
回避	4.04	1.09	.83
英語完璧主義	4.25	1.11	.87
動機づけ	3.99	1.17	.86

3.2.3 各尺度の下位尺度得点の相関係数

英語観, スピーキング不安、英語完璧主義, 動機づけの間にどのような関連性があるのかを検討するため, 各尺度の下位尺度得点間の相関分析を行った。その結果を表6に示す。特徴的な点として, 母語話者英語志向とグローバル英語志向の間にはr = .53の中程度の正の相関が見られたことから, この2つの英語志向は相反する英語観として意識されているわけではないことが示された。また, 母語話者英語志向と動機づけの間にr = .38の正の相関, グローバル英語と動機づけとの間にもr = .42の正の相関が見られ, どちらの英語志向も同程度の強さで動機づけと関連があることが示された。さらに, 英語観の下位尺度である母語話者英語志向, アイデンティティ, グローバル英語志向と英語完璧主義の間にはいずれも正の相関が見られた一方で, スピーキング不安の下位尺度である評価不安, 能力不安および回避と英語完璧主義の間にはいずれも負の相関が示された。

■表6: 下位尺度得点間の相関係数

	アイデンティティ	グローバル英語志向	評価不安	能力不安	回避	英語完璧主義	動機づけ
母語話者英語志向	.02	.53**	−.12*	.08	.14*	.38**	.38**
アイデンティティ		.23**	−.20**	−.17**	−.31**	.29**	.24**
グローバル英語志向			−.16**	.02	.08	.30**	.42**
評価不安				.53**	.52**	−.32**	−.24**
能力不安					.64**	−.26**	.10
回避						−.18**	.06
英語完璧主義							.42**

**p<.01 *p<.05

3.2.4 重回帰分析

英語観の違いがスピーキング不安, 英語完璧主義, 動機づけといった情意要因にどのような影響を与えるのかを検討するために, 母語話者英語志向, アイデンティティ, グローバル英語志向を独立変数, スピー

キング不安の3因子，英語完璧主義、動機づけをそれぞれ従属変数とする場合の重回帰分析を行った。その結果を表7〜11に示す。評価不安、能力不安，回避のいずれにおいてもアイデンティティのみが有意な負の規定要因であった。ただし，いずれにおいても決定係数（寄与率）は小さかったことから，規定要因としては弱いと考えられる。このことから，母語話者英語志向とグローバル英語志向のどちらの英語観もスピーキング不安を規定する要因にはなっていないことが示された。英語完璧主義については，母語話者英語志向とアイデンティティが有意な正の規定要因であることが示された。動機づけに関しては，英語観の3因子いずれもが有意な規定要因であった。

■表7：評価不安を従属変数とした重回帰分析結果

	偏回帰係数	標準誤差	95%信頼区間		標準偏回帰係数
			下限	上限	
母語話者英語志向	−.08	.06	−.20	.04	−.08
アイデンティティ	−.18**	.05	−.29	−.07	−.18
グローバル英語志向	−.06	.05	−.17	.05	−.07
R^2 = .06 Adj R^2 = .05					

**p＜.01 以下表11まで同様

■表8：能力不安を従属変数とした重回帰分析結果

	偏回帰係数	標準誤差	95%信頼区間		標準偏回帰係数
			下限	上限	
母語話者英語志向	.07	.06	−.06	.19	.07
アイデンティティ	−.19**	.06	−.31	−.08	−.18
グローバル英語志向	.02	.06	−.09	.14	.02
R^2 = .04 Adj R^2 = .03					

■表9：回避を従属変数とした重回帰分析結果

	偏回帰係数	標準誤差	95%信頼区間		標準偏回帰係数
			下限	上限	
母語話者英語志向	.08	.06	−.03	.20	.08
アイデンティティ	−.35**	.05	−.46	−.25	−.34
グローバル英語志向	.10	.05	−.01	.21	.11
R^2 = .13 Adj R^2 =.12					

■表10：英語完璧主義を従属変数とした重回帰分析結果

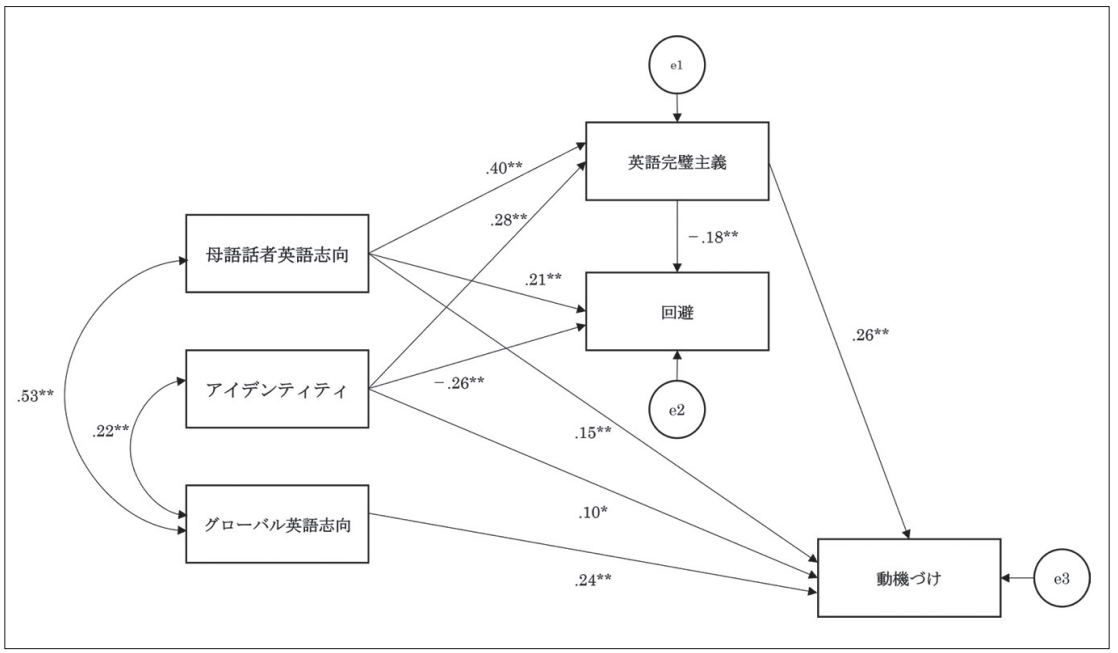
	偏回帰係数	標準誤差	95%信頼区間		標準偏回帰係数
			下限	上限	
母語話者英語志向	.35**	.06	.23	.46	.34
アイデンティティ	.28**	.05	.18	.39	.27
グローバル英語志向	.05	.05	−.05	.15	.06
R^2 = .22 Adj R^2 = .22					

■表11：動機づけを従属変数とした重回帰分析結果

	偏回帰係数	標準誤差	95%信頼区間		標準偏回帰係数
			下限	上限	
母語話者英語志向	.25**	.06	.14	.37	.24
アイデンティティ	.19**	.05	.09	.30	.17
グローバル英語志向	.25**	.05	.14	.36	.26
R^2 = .24 Adj R^2 = .23					

3.2.5 パス解析

英語観，スピーキング不安，英語完璧主義，動機づけの間の影響関係を検討するためパス解析を行った。3.2.3による相関分析および3.2.4による重回帰分析の結果を参考にして次のようにパスを想定した。(1)母語話者英語志向とグローバル英語志向の間に，アイデンティティとグローバル英語志向の間に共分散を設定した。(2)母語話者英語志向およびアイデンティティから英語完璧主義に対するパスを設定した。(3)アイデンティティから回避に対するパスを設定した。重回帰分析の結果，能力不安および評価不安については，決定率が0.10を下回っていたため除外した。(4)母語話者英語志向，アイデンティティ，グローバル英語志向から動機づけに対するパスを設定した。(5)英語に完璧さを求める傾向は英語を話すことへの不安を高めるのではないかと想定し，英語完璧主義から回避へのパスを設定した。(6)英語を専攻する学生は他専攻の学生より英語学習の重要性の認識が高い傾向があるため(吉川，2004)，英語に高いレベルを求める英語完璧主義は動機づけを高めるのではないかと想定し，英語完璧主義から動機づけへのパスを設定した。以上のパス設定によってパス解析を行った結果，RMSEA = .11と適合度が悪かった。パスはすべて有意水準5% を満たしていたが，修正指数において，母語話者英語志向から回避へのパス設定による改善度が高かったため，母語話者英語志向から回避へのパスを設定し，再度パス解析を行った。そのパス解析の結果を図1に示す。パス図に示された数字は標準化推定値である。各適合度指標の値は，χ^2 = 10.98, df = 4, p = .027, GFI = .990, AGFI = .946, CFI = .982, RMSEA = .071, SRMR = .320(注2)で，許容範囲の適合度を示した。



■図1：パス解析の結果
$\chi^2 = 10.98$, $df = 4$, $p = .027$, GFI = .990, AGFI = .946, CFI = .982, RMSEA = .071, SRMR = .320
*$p < .05$ **$p < .01$

4 考察

4.1 RQ1についての考察

母語話者英語志向はグローバル英語志向よりも有意に高い得点だったことから，高専生の英語観はグローバル英語志向よりも母語話者英語志向であることがわかった。このことは，Nakayama（2017）による英語教員養成課程の大学生と非英語教員養成課程の大学生を対象として行った比較調査で，どちらのグループの大学生も，母語話者英語志向＞グローバル英語志向＞アイデンティティの順で有意差が見られたという結果と同様であった。松浦・若生（2019）が中学生という英語学習初期の段階において学習者はすでに母語話者英語志向が形成されていることを指摘しており，高専生も高校生や大学生といった他の校種の学生と同じように，中学校で形成された母語話者英語志向の英語観を高専でも継続して持っていることがうかがえる。

しかしながら，母語話者英語とグローバル英語の関連性の捉え方には高専生の特有性が現れていると考えられる。3.2.3の相関関係の結果、母語話者英語志向とグローバル英語志向の間には相関係数$r = .53$という中程度の相関があることが示された。これは，大学生を対象とした調査を実施したNakayama（2015）では，母語話者英語志向とグローバル英語志向の間の相関係数が$r = .24$だったことと比べてより強い相関を示している。このことから，高専生は母語話者英語とグローバル英語を相反するものとしてではなく，両方を並存するものとして捉えていることが示唆された。このような高専生の英語観をもたらした理由として，まず高専生は中学校で形成された母語話者英語志向を継続したまま高専に入学してくるが，高専での学生生活の中で，アジア圏を中心とした国際交流や，将来，技術者としてアジア圏の人々と英語でやり取りをするという意識を持つことによって，グローバル英語志向が涵養されるため，両方の英語観を同時に持つようになると考えられる。そしてまた，高専生が英語学習に対して高い意欲や強い愛着を持って

いないことも関わっているのではないかと考えられる。吉川(2004)が英語専攻の学生が他学科の学生よりも母語話者英語以外の英語変種に対する受容度が低かったことを報告している。英語への関心が強く，英語を専門的に学ぶことによって，母語話者英語への規範意識や深い思い入れが培われるのであれば，技術者になることを目指し理工系専門科目を学ぶ高専生は英語力を伸ばしたいとは思っているだろうが，英語は彼らにとって主要な関心事ではないため，母語話者英語かグローバル英語かという二者択一の正しさや葛藤に悩むことがなく，両方の英語志向を共存させて受け入れているのではないかと推察される。

4.2 RQ2についての考察

パス解析の結果から，母語話者英語志向，アイデンティティ，グローバル英語志向のいずれも動機づけに対して正の影響を与えていることが明らかになったが，動機づけに及ぼすそれぞれの標準偏回帰係数はグローバル英語志向が0.24と最も高く，アイデンティティが0.10と最も低く，動機づけへの影響の強さには違いがあることが示された。グローバル英語志向は母語話者英語志向との間に中程度の相関があることから，英米の英語に限定せず母語話者英語を含めて，様々な国の英語に興味・関心を持つことが英語学習への意欲につながっているのではないかと推測される。

回避については，母語話者英語志向が正の影響を与えているのに対して，アイデンティティは負の影響を与えていることから，母語話者英語志向が強い学習者ほど英語を話すことを回避しようとする傾向がある一方で，日本人英語話者としてのアイデンティティが強い学生ほど英語を話すことにためらいや不安がないことが示された。このことから，学習者の英語を話すことへの不安を取り除くには，英語で相手と意思疎通を図るのにアメリカ人のような英語を話せなくてもかまわない，日本人英語でもコミュニケーションを取れるのだという実感を高めていくことが必要となるだろう。しかしながら，アイデンティティは英語観の3因子の中で最も有意に得点が低く，日本人英語話者としてのアイデンティティを確立するのは容易ではないことがうかがえる。アイデンティティはグローバル英語志向との間に相関関係がある一方で，母語話者英語志向との間には相関関係が見られない。そのため，日本人英語話者としてのアイデンティティを育成していくためには，日本人英語は正当な英語変種であるという理解を学習者に促進し，グローバル英語志向の英語観の涵養を図ることが有効であるだろう。

英語完璧主義については，母語話者英語志向およびアイデンティティのどちらともが正の影響を与えていることが示された。さらに英語完璧主義は動機づけに対して正の影響を与え，回避に対して負の影響を与えていることがわかった。このことから，英語完璧主義には英語運用能力を高めたいという学習者の向上意識が含まれているのではないかと推察される。母語話者英語を志向する学習者は，母語であるかのように英語を流暢に，自由自在に使いこなせることを目指し，その目標に向かって英語学習への意欲が高まるのではないかと推測される。また，英語完璧主義が回避に負の影響を与えていることから，より高い英語運用能力を目指そうとすることによって，学習者は前向きな目標意識や自己肯定感を持つことができ，英語を話すことへの不安が減少するのではないかと推察される。英語完璧主義がスピーキング回避の抑制や動機づけの向上に影響を与えるのであれば，英語完璧主義には肯定的な側面があると考えられる。そして，母語話者英語志向だけでなくアイデンティティも英語完璧主義へ正の影響を与えていることから，日本人英語話者としてのアイデンティティを持つ学習者は，流暢に英語を使いこなせる高い英語運用能力を目指していることが示唆される。日本人英語話者としてのイメージや理想像として，理解可能な英語を使用する話者としてだけでなく，英語をより思い通りに使用できる話者を想定している可能性が示唆された。このように母語話者英語志向もアイデンティティもどちらも英語完璧主義に正の影響を与えており，どちらの英語志向の学習者も高い英語運用能力を目指していることが示唆されたが，回避へは正負の符号が逆の影響を示していたことから，日本人英語話者としてのアイデンティティを持つ学習者の英語完璧主義は母語話者英語志向の英語完璧主義とは異なっていると考えられる。母語話者英語を唯一の正しいモデルとして，その英語を完璧に習得することを目指すのではなく，あくまで日本人英語話者とし

て高い英語運用能力を習得したいと捉えていることが推察される。

5 まとめと今後の展望

本研究では，高専生を対象として，彼らが有する英語観とその英語観が情意要因に及ぼす影響について検討した。その結果、以下のことが明らかとなった。(1)高専生は、母語話者英語志向＞グローバル英語志向という他の校種の学生と同様の英語観を持っている一方，母語話者英語志向とグローバル英語志向の間の相関関係が強く，両方の英語観を対立する概念とは捉えていない点に高専生の英語観の特有性が見られた。(2)母語話者英語志向，アイデンティティ，グローバル英語志向のいずれもが動機づけに対して正の影響を与えているが，最も強い影響を与えているのはグローバル英語志向であった。そして，アイデンティティはスピーキング回避に負の影響を及ぼす一方で，英語完璧主義には正の影響を与え，英語完璧主義は動機づけに正の影響を与えていることが示された。このことから，高専生の英語教育においては，グローバル英語志向と日本人英語話者としてのアイデンティティを育むことが有用であると考えられる。

高専では，英語授業の教科書に検定教科書を使用するという制限がなく，受験を意識する必要もないため，比較的自由に授業を設計することができる。各高専英語教員の創意工夫や，理工系の内容を扱った高専生のための英語教科書の開発などが行われ，高専生に適した高専独自の英語教育の在り方が模索されている。そうした中で，グローバル英語の英語概念による英語教育は高専生により適した，高専英語教育にとっての新たな道となり得る可能性を秘めていると考えられる。その第一歩として，まずは多様な英語変種を知ることが大切だと考えられる。現在多くの高専の1～3年の低学年の英語授業では高校と同様の検定教科書が使用されており，高専生が耳にする英語は教科書準拠CDによるアメリカ人を中心とした母語話者英語である。シンガポールの語学研修に行って初めてシンガポール英語を聞いたり，モンゴルやラオスからの留学生の話す英語を聞いて初めて英米以外の英語を聞いたりする学生も多くいる。そのため，まずは世界の様々な英語の音声や語彙，表現に触れさせ，興味・関心を持たせることがグローバル英語志向を涵養する入口となるだろう。

また，これまでの高専生の動機づけについての知見によって，高専生は内発的動機づけが弱く，外発的動機づけが特徴であることが指摘されている（石川，2012; 倉増，2020, 2023）。高専生には，英語は何等かの目標を達成するための道具であると捉える言語道具主義の傾向がみられる。英語をコミュニケーションのための道具と捉えることは，母語話者英語への愛着やこだわりが薄いという点では有用な面もある。しかしながら，英語を単なる道具と捉えて，英語が使いこなせればそれでかまわないと思うことによって，言語には文化的背景や社会的背景があることに思い至らず，対話する相手の文化への理解や配慮を軽視する懸念がある。グローバルに活躍できる技術者へ育成するためには，多様な人々と対等なコミュニケーションを図れる態度を涵養することが大事だろう。

なお本研究では，高専生の英語観の時間的変容を捉えることにまで踏み込むことができなかった。高専は5年間の一貫教育のため1年生から5年生までの間，一貫して同じ英語観を持っているとは考えづらい。縦断的な調査を行うことで，中学校で身についた英語観は高専に入ってからの英語観にどの程度影響しているのか，そうであればそれはどれくらいの期間に渡って影響し続けるのか，英語観の変容に最も影響を与える要因は何なのか（授業、国際交流、海外研修など），5年生になり就職を意識することで英語観に変化が起きるのかといった点を明らかにすることができるだろう。こうした縦断的な検討によって，高専生の英語観を長期的な観点でより深く理解して，グローバル英語の概念を教育実践に活かしていく必要があると考えられる。

注

1. 石原（2020）では，「評価不安」の項目が13, 7, 10, 2, 15, 4, 1, 5,「能力不安」の項目が14, 8, 11, 16, 12,「回避」

の項目が3, 6, 9であった。石原(2020)のスピーキング不安尺度を使用した長友(2022)では,「評価不安」の項目が1, 7, 5, 13, 10, 4,「能力不安」の項目が16, 11, 14, 8, 15, 2,「回避」の項目が3, 6, 9, 12であった。各因子に属する質問項目には若干の違いがみられるが, スピーキング不安尺度の構成要素と同様の3因子構造の結果となっている。

2. 適合度指標としてSRMR(standardized root mean square residual)を用いることの有用性については, In'nami and Koizumi(2011)を参照。

謝辞

本研究を発表する貴重な機会を与えてくださった公益財団法人 日本英語検定協会と関係者の皆様、選考委員の先生方に深く御礼申し上げます。特に論文を作成するにあたり, 助言者としてご助言くださいました斉田智里先生に深く感謝いたします。また, 調査にご協力いただいた国立高等専門学校機構広島商船高等専門学校の学生の皆様に御礼申し上げます。

引用文献

Chiba, R., Matsuura, H., & Yamamoto, A. (1995). Japanese attitudes toward English accents. *World Englishes*, 14, 77-86.

Dewaele, J.-M. (2017). Are perfectionists more anxious foreign language learners and users? In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), New Insight into language anxiety: Theory, research and educational implications (pp. 70-90). Multilingual Matters.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. Longman.

Gregersen, T. & Horwitz, E. K. (2002) Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal* 86(4), 562-570.

花元宏城. (2010).「英語変種に対する日本人大学生の言語態度について:matched-guise技法を用いた調査」.『アジア英語研究』12, 21-37.

Horwitz, E. K. (1986) Preliminary evidence for the reliability and validity of a foreign language anxiety scale *TESOL Quarterly* 20(3), 559-562.

In'nami, Y., & Koizumi R. (2011). Structural equation modeling in language testing and learning research: A review. Language Assessment Quartely, 8(3), 250-276.

一般社団法人国際ビジネスコミュニケーション協会 (2016).『TOEIC® Program DATA & ANALYSIS 2016年度受験者数と平均スコア』

一般社団法人国際ビジネスコミュニケーション協会 (2017).『TOEIC® Program DATA & ANALYSIS 2017年度受験者数と平均スコア』

一般社団法人国際ビジネスコミュニケーション協会 (2018).『TOEIC® Program DATA & ANALYSIS 2018年度受験者数と平均スコア』

一般社団法人国際ビジネスコミュニケーション協会 (2019).『TOEIC® Program DATA & ANALYSIS 2019年度受験者数と平均スコア』

一般社団法人国際ビジネスコミュニケーション協会 (2020).『TOEIC® Program DATA & ANALYSIS 2020年度受験者数と平均スコア』

一般社団法人国際ビジネスコミュニケーション協会 (2021).『TOEIC® Program DATA & ANALYSIS 2021年度受験者数と平均スコア』

一般社団法人国際ビジネスコミュニケーション協会 (2022).『TOEIC® Program DATA & ANALYSIS 2022年度受験者数と平均スコア』

石原知英 (2020).「英語スピーキング不安の構造とその低減のための取り組み:ペアトークとグループプレゼンテーション」.『九州英語教育学会紀要』48, 39-48.

石川希美. (2012).「英語学習状況に関する調査ー高専生の英語嫌いと学習意欲ー」.『全国高等専門学校英語教育学会研究論集』29, 41-50.

磯田貴道 (2008).「英語スピーキング抵抗感尺度の作成」『広島外国語教育研究』11, 41-49.

Jenkins, J. (2000). *Phonology of English as an International Language.* Oxford University Press.

Kachru, B. B. 1985. (1985). Standards, codification, and sociolinguistic realism: The English language in the outer circle. In R. Quirk, & H. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge University Press.

川又正之. (2007).「非母語話者の英語 (3)ー改訂版中学校英語教科書における『変種』の取り扱いについて」『外国語教育論集』29, 103-116.

Kawashima, T. (2018). A longitudinal study of speakers on CDs for senior high school English textbooks in Japan. *Asian English Studies*, 20, 26-46.

倉増泰弘. (2020).「高等専門学校生の英語学習に対する動機づけを測定する尺度」『中国地区英語教育学会誌』50, 1-10.

倉増泰弘. (2023).「高専生の英語学習動機の変化に関する考察」『全国高等専門学校英語教育学会研究論集』42, 51-60.

Matsuda, A. (2003). The ownership of English in Japanese secondary schools. *World Englishes, 22*(4), 483-496.

Matsuura, H., Chiba, R., & Yamamoto, A. (1994). Japanese college students' attitudes towards non-native varieties of English. In D. Graddol & J. Swan (Eds.), *Evaluating Language* (pp. 52-61). Multilingual Matters.

引用文献

松浦浩子・若生深雪. (2019).「多様な英語発音に対する日本人学習者の態度－中・大比較から見えること－」.『商学論集』87(4), 1-13.

Mimatsu, T. (2011). ELF versus FFL: Teaching English for international understanding in Japan. In A. Archibald, A. Cogo & J. Jenkins (Eds.), *Latest trends in ELF research* (pp. 251-267). Cambridge Scholars Publishing.

Munezane, Y. (2013). Motivation, ideal L2 self and valuing of global English. In M. T. Apple, D. Da Silva, & T. Fellner (Eds.), Language learning motivation in Japan (pp. 152-168). Multilingual Matters

長友隆志. (2022).「英語スキット学習を通したスピーキング不安の改善」『外国語教育メディア学会九州・沖縄支部』22, 17-25.

中井英民. (2009).「EIL（国際語としての英語）をめぐる議論と英語教育―英語（音声）学習モデルに関する中・高教員とALTの意識をもとに」『外国語教育』35, 1-39.

Nakamura, Y., Lee, J. S. & Lee, K. (2018). English as an International Language perception scale: Development, validation, and application. *Language, culture and communication, 50*, 189-208.

Nakayama, T. (2015). An EIL awareness questionnaire: Development and Validity in a Japanese context. *The Faculty of Letters of Jissen Women's University annual reports of studies,* 57, 11-19.

Nakayama, T. (2017). EIL awareness among Japanese pre-service EFL teachers and EFL students: Who are their ideal role models? *Jissen Women's University CLEIP Journal,* 3, 35-41.

大石海. (2022).「小学校英語教科書の登場人物と英語音声の批判的分析－表面的な多様性と母語話者志向－」『東京大学大学院教育学研究科紀要』91, 109-118.

Oxford, R. L. (1999). Anxiety and the language learner: New insights. In J. Arnold (Ed.), Affect in language learning (pp.55-67). New York: Cambridge University Press.

Price, M. L. (1991). The subjective experience of foreign language anxiety. Interviews with highly anxious students. In E. K. Horwitz & D. J. Young (Eds.), Language anxiety: From theory and research in classroom implications (pp. 101-108). Upper Saddle River, NJ: Prentice Hall.

Saito, A. (2014). Is English a nuisance or an asset? Japanese youth's discursive constructions of language attitudes. *System,* 44, 13-23.

関谷弘毅. (2009).「スピーキング学習における英語学習観尺度と性格・情緒尺度の開発」『東京大学大学院教育学研究科紀要』48, 147-154.

Smith, L. (1983). *Readings in English as an International Language.* Pergamon.

Taguchi, T., Magid, M. & Papi, M. (2009). The L2 motivational self system among Japanese, Chinese and Iranian learners of English: A comparative study. In Z. Dörnyei, & E. Ushioda (Eds), *Motivation, language identity and the L2 self* (pp. 66-97). Multilingual Matters.

Takahashi, C. (2021). Attitudes toward the global spread of English and motivation to study language: A graded response model analysis. Second Language Studies, 39, 77-113.

Tokumoto, M. & Shibata, M. (2011). Asian varieties of English: Attitudes towards pronunciation. *World Englishes,* 30, 392-408.

行森まさみ. (2015).「日本人英語教師の英語観：「国際語としての英語」を中心として」『「英検」研究助成報告』27, 175-183.

資料1：英語観に関するアンケート

【1】英語の発音について，あなたが思うことについてお聞きします。

選択肢からあてはまる数字を選んでください。

6. とてもそう思う
5. そう思う
4. 少しそう思う
3. あまりそう思わない
2. そう思わない
1. 全くそう思わない

(1) 授業では，アメリカ人のような英語を母語とする人の発音による英語教材を使うべきだ。

(2) 日本の学校では，日本人らしい英語の発音を教えるべきだ。

(3) リスニングでは，アメリカだけでなく様々な国の人の発音による英語教材を使うべきだ。

(4) 日本人はできるだけアメリカ人のような英語を母語とする人の英語の発音に近づくように努力するべきだ。

(5)「君の英語の発音はアメリカ人みたいだね」と英語を母語とする人からほめられたらうれしい。

(6) リスニングでは，英語力のある日本人の発音による英語教材を使うべきだ。

(7) 授業では，アメリカだけでなく様々な国の人の発音による英語教材を使うべきだ。

(8) 日本人らしい英語の発音ができる人を尊敬すべきだ。

(9) 私はアメリカ人のような英語を母語とする人の発音を身につけたい。

(10) 自分の英語が日本人なまりの発音だと言われても気にならない。

(11) インド英語やシンガポール英語など様々な国の人の英語の発音を理解できるべきだ。

資料2：スピーキング不安に関するアンケート

【2】英語を話すことについて，あなたが思うことについてお聞きします。

選択肢からあてはまる数字を選んでください。

6. とてもよくあてはまる
5. あてはまる
4. 少しあてはまる
3. あまりあてはまらない
2. あてはまらない
1. 全くあてはまらない

(1) 私は，クラスの人に英語の発音が下手だなと思われないか心配です。

(2) 私は，先生に英語のスピーキング力を評価されているとき、不安になります。

(3) 私は，できれば人と英語で話したくありません。

(4) 私は，クラスの人の前で英語を話すとき、緊張します。

(5) 私は，私の話す英語が先生に上手く伝わっているかどうか心配です。

(6) 私は，人と英語で話すことは避けたいです。

(7) 私は，クラスの人に英語を話すのが下手だなと思われないか心配です。

(8) 私は，英語を話す力がないと思います。

(9) 私は，英語で話さなければならないときでも，日本語で話したいと思うことがよくあります。

(10) 私は，先生が私の英語の間違いをチェックしているときに不安になります。

(11) 私は，今の英語力では英語で話すことはできないと思います。

(12) 私は，英語で話さなければならないときはできるだけしゃべらないようにしています。

(13) 私は，私の話す英語がクラスの人に上手く伝わっているかどうか心配です。

(14) 私は，自分が考えていることを英語で相手に伝えることができないと思います。

(15) 私は，クラスの人が英語を上手に話しているときに不安になります。

(16) 私が話す英語は，相手に意味が伝わらないと思います。

資料3：英語完璧主義に関するアンケート

【3】英語ができるようになることについて，あなたが思うことについてお聞きします。

選択肢からあてはまる数字を選んでください。

6. とてもよくあてはまる
5. あてはまる
4. 少しあてはまる
3. あまりあてはまらない
2. あてはまらない
1. 全くあてはまらない

(1) 英語ができるようになるということは，どんな英語でも正確に日本語に訳せるようになることだ。

(2) 英語ができるようになるということは，文法が完璧にわかるようになることだ。

(3) 英語ができるようになるということは，どんな英語でも聞いたり読んだりして理解できるようになることだ。

(4) 英語ができるようになるということは，言いたいことが何でも話せるようになることだ。

(5) 英語ができるようになるということは，アメリカ人やイギリス人のように話せるようになることだ。

(6) 英語ができるようになるということは，学校のテストで高得点が取れるようになることだ。

(7) 英語ができるようになるということは，検定試験(英検など)でよい結果が出せるようになることだ。

資料4：動機づけに関するアンケート

【4】英語を学ぶことについて，あなたが思うことについてお聞きします。

選択肢からあてはまる数字を選んでください。

6. とてもよくあてはまる
5. あてはまる
4. 少しあてはまる
3. あまりあてはまらない
2. あてはまらない
1. 全くあてはまらない

(1) 今後、英語の授業が学校やその他の場所で開かれたら，受けてみたい。

(2) 私は英語学習を一生懸命にやっている。

(3) 私は英語学習に努力を惜しまない。

(4) 私は英語学習を頑張っていると思う。

今回の報告について

～ 日本英語検定協会 会長･名誉顧問より ～

名誉顧問　長　勝彦

研究助成選考委員の厳しい審査をパスした各報告書は，いずれも研究計画，研究結果ともに，レベルの高い素晴らしい報告書であった。特に実践部門においては，将来，これらの研究成果を公開授業，たとえば，全国英語教育研究団体連合会（通称：全英連）などを通して，多くの英語教師との共有を目指して下さい。素晴らしい成果が全国に広まって行くことを期待しております。

◆　◆　◆

名誉顧問　小池　生夫

今回の応募論文は研究に関する論文が４件，実践に関するものが５件，調査部門から１件あり，計10論文が審査を通った。この中で集中しているのは研究と実践部門で，全体で10編中９編になり，調査部門は1編である。この傾向は昨年も同じであり，ここ数年来の傾向となっている。助言指導の先生方もメンバーが交代して若返り，一層熱心に指導してくださるのが論文を読むと見えるのは嬉しい限りである。

今回の研究成果は従来の評価に関する分野での新しい英検の研究と実践による莫大な資料の利用に基づく歴史的発展によるものでもあるが，国際的に見て米国で始まった独創的理論研究が，わが国の英語テスト理論の研究に一層直接の利用を与え，本研究の研究部門や実践部門と関係があるペーパーの作成に関連して，統計学研究においても長足の進歩を見せるに至った。こういう研究の継続，発展に依る結束，協力しての活動による貢献度は大きい。現在ではその流れは多角的に拡大，進歩し，テスト研究の発展の速さも幅も広がってきたという背景がある。最近では，この流れに本格的な興味を持つ若手の研究者が参加し，増えてきて，競って研究をするという傾向が出てきている。富士山型の応用言語学的研究が末広がりに発展するようになってきた。また国内から海外の日本人の情報交換の拡大を含めた国際的刺激を受けて，発展してもいる。

◆　◆　◆

名誉顧問　村木　英治

コロナ禍がまだ完全に過去のものとなっていない現在，しかし，一方で，以前の生活が確実に戻ってきている。研究活動は研究トピックの選択から開始して，このEIKEN BULLETINのような研究専門誌での掲載という結実まで，通常，複数年に渡るものである。したがって，今回の掲載論文のほとんどは，コロナ禍というトンネルを潜り抜けてきたのだろう。AIという最新のトピックもあったが，ネットでの動画映像のテーマ選択など，その影響は明らかだ。日本人の英語教育に関連する研究は時代の要請に応えるものが主だから，そうであることは健全な傾向だといえる。コロナ禍に負けず，時代の変化に対応できるエネルギーの持続が感じられた。また，特にどの論文も添付資料が充実している。個人的には，三上氏の論文が，時宜を得て，これからの英語教育における必要性と可能性によく応えているものだと思った。

会長　吉田　研作

今回の多くの研究は，実際の学習状況の中から生まれた課題を基にしているように思う。昔の研究は，理論が先行し，その理論に基づいた仮説を検証するために，恣意的に被験者を選び，さらにその被験者を実験群と統制群に分けて実験をした。その頃は，できるだけ純粋な結果を得るために，実際の教室ではなく，実験室のような場所で行われた。そのため，出てきた結果は，実験室のような無菌状態では有効でも，普段の教室では，なかなか確認できなかった。それに対して，現在の研究は，実際の学習状況の中から課題を発見し，それがどのような要因によるものかを仮説として立て，検証するという方法が取られる。今回は，まさにそのような現実をベースにした研究で，実際に役立つものが多いと思う。

◆　◆　◆

名誉顧問　和田　稔

第35回「英検」研究助成の入選者の論文を読んだ。論文は研究助成の選考委員の厳しい審査をパスしたもので，それらの論文を読むことは楽しく，また，知的な関心を刺激される。

私は第35回のEIKEN BULLETINを一つの興味・関心をもって読み始めた。それは論文が日本の小・中・高等学校における英語教育の実際を直接的に視野に入れているのかという興味・関心である。研究論文をそのような観点から読むことは学術的論文の普遍性を歪めることであろう。しかし，日本の学校英語教育における実際や課題など直接的に研究し，その結果を発表するという問題意識もまた忘れてはならないのではないだろか。

例えば，このような観点から論文を読むと小学校における英語教育を直接的に研究対象にした論文が無いことに気付く。日本の小学校での英語教育は長い年月にわたっての紆余曲折を経て教科に位置づけられたことは周知の事実であろう。小学校における英語の指導や学習にはいろいろな課題があるはずである。

ISSN 1348-7949

EIKEN BULLETIN vol. 35 2023 非売品

2024年2月2日　初版 発行

編集・発行　公益財団法人 日本英語検定協会

〒162-8055　東京都新宿区横寺町 55
https://www.eiken.or.jp
印刷・製本／日新印刷株式会社

無断複製・転載を禁じます。　©2024 公益財団法人 日本英語検定協会