

日本人英語学習者のスピーキング vs. ライティングパフォーマンスの比較分析のための指標

—学習者コーパスに基づくアプローチ—

高知県／土佐女子中学高等学校 講師 野村 真理子

申請時：東京都／東京外国語大学大学院在籍

概要

本研究は、日本の中学・高等学校の外国語科の新学習指導要領を踏まえ、話す活動と書く活動を組み合わせて表現力を育成する効果的な指導を模索するための基礎調査として、中学生・高校生の日本人英語学習者からスピーキングとライティング両方の産出データを収集し、学習者コーパスを構築することにより、同一学習者集団のスピーキング vs. ライティングパフォーマンスを比較分析し、比較測定のために有効な指標を調べたものである。分析に使用したコーパスデータは、中学3年生～高校3年生324名から収集した同一学習者が同一テーマで産出したスピーキングとライティングのデータである。使用語彙、10の言語特徴、5つの言語項目のエラーおよび正用率について、話し言葉 vs. 書き言葉で比較分析を行った結果、調査した多くの項目がモードの違いによる差異を示し、日本人中高生の産出モードの異なるパフォーマンスデータを比較分析するのに有効な指標が明らかになった。

1 はじめに

外国語科の新中学校学習指導要領（平成20年3月公示）および新高等学校学習指導要領（平成21年3月公示）では、これまでの「聞くこと」、「話すこと」に加え、「読むこと」、「書くこと」のコミュニケーション能力を養うことが明示され、4技能を総合的に育成する指導、4つの領域の言語活動の統合に重点が置かれている。産出面の改訂事項を見ると、中学校では、「話すこと」の領域に、「与えられたテーマについて簡単なスピーチをすること」、「書

くこと」の領域に、「語と語のつながりなどに注意して正しく文を書くこと」と「自分の考えや気持ちなどが読み手に正しく伝わるように、文と文のつながりなどに注意して文章を書くこと」という指導事項が新たに示された。高等学校では科目構成が変更され、話したり書いたりする言語活動を中心に、情報や考えを伝える能力の向上を図るための科目として「英語表現」が創設された。「英語表現Ⅰ」には、「与えられた話題について、即興で話す」という言語活動が含まれているが、これは、これから話す内容を前もって準備せずに即興で事実や意見、感情などを伝えることができるようになることが必要であるという観点から、言語活動に含まれたものである。

新学習指導要領においては、技能間の有機的なつながりが重視されているので、話すことと書くことが一連となった活動も有効であると考えられる。この場合、「話す→書く」、「書く→話す」という活動の組み合わせが考えられるが、即興で話す活動が重視されていることを踏まえ、まずある話題について即興で話す活動をした後、同じ話題について書く活動をするのも1つの効果的な指導であると考えられる。このような指導を実施するにあたって、中高生のスピーキングとライティングパフォーマンスを調査し、話し言葉と書き言葉の習得の実態を把握しておくことは有益であると思われる。

中高生の書き言葉については、中高生1万人の英作文を収集して電子化した大規模な学習者コーパス—Japanese EFL Learner (JEFL) Corpus（投野, 2007）の構築、公開などにより、中高生が書く英文の実態解明が進みつつある。

一方、話し言葉の方は、日本人のインタビューテ

ストにおける発話を収集した大規模なスピーキングコーパス The NICT JLE Corpus (和泉・内元・井佐原, 2004) が公開され、研究が進みつつあるが、これは主に成人の学習者を対象としたものであり、中高生の実態を把握することは難しい。和泉他で用いられているのと同じインタビューテストのタスク(一部修正)を中高生に実施して、中高生のスピーキング能力の実態解明を試みた野田(2008)のような研究もあるが、中高生が産出した発話を収集して、話し言葉の実態を解明しようとする研究は、実施上の困難さもあり、まだ極めて限られているのが現状である。

さらに、中高生と同じ学習者が産出した話し言葉と書き言葉を収集した研究はほとんどなく、話し言葉 vs. 書き言葉の実態や発達過程はまだほとんど解明されていない。

本研究は、話す→書くという活動の組み合わせによる効果的な指導のための基礎調査として、中学・高等学校の日本人英語学習者からスピーキングとライティング両方の産出データを収集し、学習者コーパスを構築することにより、同一学習者集団のスピーキング vs. ライティングパフォーマンスにおける言語特徴を比較分析し、比較測定のための指標となりうるものを明らかにしようとする試みである。

2 先行研究

2.1 中間言語の研究

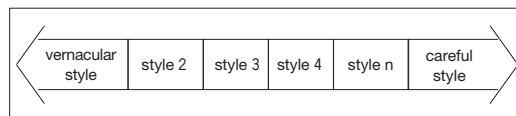
第二言語の学習者の言語は、構造的に母語と目標言語の中間に位置する学習者独自の体系であるという意味で、Selinker (1972) は、このような学習者言語を中間言語 (interlanguage) と呼んだ。中間言語には、学習者に共通の誤用や一定の習得順序などの共通性が見られ、その発達段階に規則性が見られるが、同時に学習者の言語には変異性が見られる。中間言語の変異性については、これまで数多くの実証的研究が行われ、変異に影響するさまざまな要因が指摘されている (Ellis, 1994)。

Larsen-Freeman (1975) や Tarone (1985) は、タスクの違いにより文法項目の正確度が推移することを示した。Larsen-Freemanでは、5つの異なるタスク (reading, writing, listening, imitating, speaking)、Tarone では、4つの異なるタスク (written

grammaticality judgment task, oral description, oral narration, oral interview) が用いられた。これらの研究のタスクには、話すタスクと書くタスクが含まれており、話し言葉 vs. 書き言葉というモードの違いにより学習者のパフォーマンスが異なることが示唆された。Vann (1979) でも、モードの違いにより学習者の表現方法に差異があることが示された。さらに、タスクの planning time の違い、学習者の第二言語の熟達度レベル、同じテーマのタスクを一定期間後に実施した場合の practice effect などによっても言語項目の正確度が推移することが実証されている (Ellis, 1987; Young, 1988)。

Tarone (1983: 152) は、図1のように、中間言語を言語形式に向けられる注意の程度によって定義されるスタイルの連続体ととらえ、注意がほとんど払われない日常的なスタイルである vernacular style と最も注意が行き届いたスタイルである careful style を両極端とする連続体内で、中間言語は変化するという仮説を唱えた。

▶ 図1：中間言語の連続体



この仮説に基づき、Tarone (1985) における文法項目の正確度は、タスクにより必要とされる言語形式への注意の度合いにより推移しているとした。また、Tarone (1985) では、学習者が自分の産出したものをチェックし、誤っていると考えられるとき訂正しようとする monitoring が、3人称単数現在 (-s) などの正確度の向上に影響していること、母語の違い (日本語とアラビア語) により同じ文法項目で異なる正確度を示していることが述べられており、monitoring や母語も中間言語変異の要因であることが実証されている。Tarone (1982) は、monitoring は学習した目標言語の規則体系によって、careful style に働くと見なしている。また、monitoring が働くためには、訂正のために十分な時間がなければならないとされており、monitoring は planning time の有無ともかわってくる。

以上のように、中間言語の変異性を左右する因子はいろいろあるが、本研究では、日本人中高生への話す→書くの指導を視野に入れ、話し言葉 vs. 書き

言葉というモードの違いおよび学習者の熟達度レベルの違いに焦点を当てる。

2.2 話し言葉 vs. 書き言葉の比較研究

1990年代より、学習者が話したり、書いたりした英語を収集して電子化した学習者コーパスを用いた研究が行われるようになり、話し言葉 vs. 書き言葉を比較した注目すべき研究がいくつか見られる。

Haynes (1992) は、3つの熟達度レベルから成る学習者64人の話し言葉と書き言葉データを同一テーマで収集して、Biber (1986) の multi-dimensional approach を用いて、11個の言語特徴を分析し、話し言葉と書き言葉の変異を調査している。この研究では、テーマのタイプは叙述文 (narrative) で、“What was your most frightening experience?” というテーマについて、最初に発話が収集され、2、3週間後に作文が収集されている。dimension がモード差を表し、言語特徴が発達指標として作用するという結果を示している点で注目すべき研究である。

日本人英語学習者の研究では、The NICT JLE Corpus と JEFLL Corpus を用いた話し言葉 vs. 書き言葉の比較分析があり (Abe, 2007; 小林, 2008; 小林・山田, 2008; 鈴木・投野, 2008; 投野, 2008; 三浦, 2008)、モード差および学習段階により差を示す項目が示されている。

小林 (2008) は、The NICT JLE Corpus と JEFLL Corpus の高頻度語100タイプの頻度パターンにコレスポンデンス分析を実行し、モード間に顕著な対比が存在し、高頻度語100タイプの頻度分布と習熟度レベルの間に相関関係が見られることを明らかにしている。現在形の一般動詞、過去形の動詞、1人称の代名詞、従属接続詞が書き言葉を特徴づけるカテゴリー、現在形の be 動詞、前置詞が話し言葉を特徴づけるカテゴリーとなっている。上記の研究はすべて、異なる学習者集団、異なるタスクのデータを比較したものである。

話し言葉と書き言葉の産出モードの違いによる言語特徴を比較するには、投野 (2008: 132) が述べているように、「厳密には同一タスクで同一被験者の行う言語使用データを採取する必要がある」が、日本人の同一学習者の話し言葉と書き言葉を比較した研究はほとんどない。同一タスク (1枚の絵の描写) で同じ大学生31人の発話と作文を収集している Abe (2003)、同一テーマで中高生40人の発話と作

文を収集している野村 (2009) のように、これまでの研究はまだデータ数が少ないが、産出モードの違いにより学習者のアウトプットに違いが見られたことが示されている。

野村 (2009) では、中学3年生20名、高校3年生20名が産出した同一テーマでの発話と作文のコーパスデータを用いて、産出モードの違いによる言語特徴を単語および品詞の3連鎖 (trigram) の比較から分析した。単語連鎖の分析では、発話の特徴として、(a) filler, 自己訂正, 繰り返しを含む連鎖, (b) be 動詞, very を含む連鎖, (c) 短縮形の使用のような連鎖が見られ、作文では、(a) that 節, (b) if 節, (c) 不定詞構文の連鎖が特徴的で、構造が複雑化し、複文的になっていることがわかった。品詞連鎖においては、発話では filler を含む連鎖、代名詞の繰り返しを含む連鎖が特徴的で、作文では、前置詞+名詞の連鎖、冠詞、法助動詞の挿入傾向、動詞+形容詞の連鎖、文頭における代名詞の使用が特徴的であった。

同一学習者のデータではないが、同様のタスクを用いて産出モードの違いによる言語特徴を比較した研究に Abe (2003)、朝尾 (2008) がある。Abe は、1枚の絵の描写のタスクで発話 vs. 作文コーパスにおけるエラーを調べ、エラー分析から産出モードの違いによる中間言語変異を調べたもの、朝尾は、Traveling という同一テーマで大学生が産出した発話 vs. 作文コーパスで、語彙頻度、単語の3連鎖、接続詞の使用などの比較をしたものである。

これまでの日本人英語学習者のコーパスに基づく研究において、発達指標になりうるものとしては、名詞句、代名詞、動詞、冠詞、前置詞、単語・品詞連鎖、談話標識、語彙指標などが示唆されており、これらの項目の多くはモード差も観察されている。語彙指標の観点からの調査では、総語数が発話と作文それぞれの習熟度を判別することが統計的に実証されているが (水本, 2008; 小林, 2009)、発話 vs. 作文のモード差を示す指標になるかはまだ検証されていない。日本人中高生の書き言葉は、JEFLL Corpus の公開などにより、大規模なデータを利用することができるが、話し言葉の方はデータ収集や実態解明が遅れている。5つのスピーキングタスクで中高生225人から発話データを収集し、いくつかの語彙指標の妥当性を調査している小泉 (2007) のように、ある程度の規模のデータを収集し、研究することが必要である。

英語学習者の話し言葉 vs. 書き言葉の言語特徴を分析するにあたって、英語母語話者の話し言葉 vs. 書き言葉の特徴を把握しておくことも必要である。小規模なデータに基づいたものではあるが、話し言葉と書き言葉の特徴を調べた注目すべき研究がいくつかある (Chafe, 1981; Ochs, 1979; Tannen, 1982)。

Chafe (1981) は、14人の食卓での会話 vs. 学術論文を比較したもので、スピーキングとライティングでは両極端のスタイルに位置するもの (自然な会話 vs. 形式的な学術論文) であり、したがって、言語特徴の分析結果は、最大限のモード差を示していると考えられる研究である。話し言葉 vs. 書き言葉の特徴として、fragmented vs. integrated quality と involvement vs. detachment を挙げ、それぞれの特徴を示す項目について、モードの違いによりどれくらいの頻度差があったかを示している。

話し言葉の断片的な性質として、アイデアユニットの最初に and, but, so, because を用いる傾向は話し言葉と書き言葉を比較すると、話し言葉の方が約4倍頻度が高くなっている。書き言葉の integration の性質を示す言語特徴として、名詞化、分詞 (現在分詞および過去分詞)、属性の形容詞、前置詞句、補文 (that ~, to ~)、関係代名詞などが挙げられている。これらの項目は、書き言葉の方で、約2倍~11倍多く用いられている。

また、話し言葉の involvement を示すものとして1人称代名詞の使用 (話し言葉での使用頻度が約12倍)、書き言葉の detachment を示すものとして受動態の使用 (書き言葉での使用頻度が約5倍) が挙げられている。

同じ話者が同じテーマについて話す→書くの順で産出したデータを収集して、話し言葉 vs. 書き言葉の言語特徴を比較分析した試みが、Ochs (1979) と Tannen (1982) で見られる。Tannen のデータは、上で述べた Chafe (1981) の言語特徴を反映したデータになっているが、書いた英文の方が話した英文よりもずっと短く (例: 340語 vs. 76語) になっており、話し言葉 vs. 書き言葉の fragmented vs. integrated という差が顕著に現れていることが特徴である。

近年では、英語母語話者の話し言葉と書き言葉をコーパス化した大規模なデータ (Longman Spoken and Written English Corpus) を用いて、会話、小説、ニュース、学術論文の4つのレジスターにおけるさ

まざまな文法項目の使用頻度の差を調査した研究がある (Biber, Johansson, Leech, Conrad & Finegan, 1999)。上で述べた3つの母語話者の研究で示されているモードの違いによる言語特徴をこの Biber et al. で検証すると、ほぼ一致した傾向が見られた。表1は、英語母語話者の先行研究で調査されているモードの違いによる言語特徴で、日本人英語学習者のデータでもこれまでに調査されているもの、また今後調査する価値があると思われるものについて、話し言葉 vs. 書き言葉における使用頻度の差をまとめたものである。○がついている方が使用頻度が高いことを示している。なお、頻度差は必ずしも統計的に有意な差ではない。

■ 表1: 話し言葉 (S) vs. 書き言葉 (W) の言語特徴

言語特徴	S	W
conjunctions: and, but, so, because	○	
first person references	○	
articles		○
attribute adjectives		○
N + of		○
prepositions		○
subordinate clauses		○
relative clauses		○
complement clauses: that and to		○
active voice	○	
passive voice		○
present tense	○	
past tense		○

母語話者データでは、用いたデータが異なっても産出モードの違いによる言語特徴には一定の傾向が見られるが、これまでの日本人英語学習者の先行研究では、使用したコーパスデータにより、表1の項目について異なる結果が示されているもの (1人称の代名詞や前置詞など) があり、今後さらなる検証が必要である。

3 研究の目的

本研究の目的は、日本の中学・高等学校レベルの同一学習者集団のスピーキング vs. ライティングパフォーマンスをコーパス言語学的手法を用いて比較

分析し、日本人中高生の話し言葉 vs. 書き言葉の特徴づける指標を明らかにすることである。本研究では、母語話者および英語学習者の先行研究でモードの違いによる言語特徴として指標となりうるものが示唆されている言語項目に焦点を当て、それらが中学・高等学校レベルの日本人英語学習者のパフォーマンスで指標として有効であるかを調べることに焦点を当てる。そして、明らかになった指標に基づき、中高生のスピーキング vs. ライティングパフォーマンスの比較測定について示唆を与えたい。

4 研究方法

4.1 データ収集

4.1.1 参加者

参加者は、私立と県立合わせて4校の中学3年生～高校3年生で、同一テーマでスピーキングとライティング両方のタスクを行った324名をコーパスデータとした。英語の能力に関して、学習者のレベルは、全体としては、低 (L)、中 (M)、高 (H) のすべてをカバーしているが、表2に示したように、学年によりレベルに偏りがある。表2の英検取得者の人数は、参加者の自己申告である。なお、高2には1名準1級合格者が含まれている。参加者が多かった高1の学習者集団は、英語のレベルは多様で、ほぼすべてのレベルをカバーしている高1の母集団であると考えられる。参加者の約半数は英語科および国際に関する学科に属し、比較的英語が好きな学習者であり、日常的に英語を話す活動もより多くしており、ある程度の量のまとまった英文を書く訓練を受けているクラスの学習者も含まれている。一方でその他の普通科の参加者には、英語を話したり、

■ 表2：参加者の概要

	中3	高1	高2	高3	計
レベル	M, H	L, M, H	H	M, H	L, M, H
参加者数	24	221	38	41	324
英検3級取得者	24	46	1	6	77
英検準2級取得者	0	17	30	15	62
英検2級取得者	0	0	4	13	17

まとまった英文を書いたりする練習を日頃はほとんどしていない学習者も多く含まれている。

4.1.2 収集方法

まずスピーキングのタスクを実施し、その後2週間以内くらいをめどに、スピーキングと同一テーマによるライティングのタスクを実施した。本研究では、話す→書くという順でタスクを行った場合の中間言語の特徴に重点を置き、できるだけ自然な発話の実態を把握するため、先にスピーキングを実施し、発話の内容をまだ記憶している時期に、主な発話内容は変えないようにして作文を書かせた。後で同一テーマで作文を書くことは生徒には前もって知らせていない。

スピーキングとライティングのテーマは、英検3級～2級の合格者の実際の英語使用に対する自信の度合いを示した Can-do リスト ((財) 日本英語検定協会, 2006) の中で、「話す」、「書く」に共通している表3のテーマを参照して下記の A, B, C に決定した。

テーマ：A「ご飯派それともパン派？」

B「将来訪れたい国 (または日本の地域)」

C「印象に残った学校行事」

■ 表3：Can-do リスト

3級	話す	物ごとの「好き」「嫌い」とその理由を簡単に述べるができる。(動物, 食べ物, スポーツなど)
	書く	物ごとの「好き」「嫌い」とその理由を書くことができる。(食べ物, スポーツ, 音楽など)
準2級	話す	自分の将来の夢や希望について、話すことができる。(訪れたい国, やりたい仕事など)
	書く	自分の将来の夢や希望について、書くことができる。(訪れたい国, やりたい仕事など)
2級	話す	印象に残った出来事について、話すことができる。(旅行, イベントなど)
	書く	印象に残った出来事について、その内容を伝える文章を書くことができる。(学校行事, 旅行など)

タスクは以下のとおりである。タスクの制限時間については、パイロット研究として行った先行研究(野村, 2009)で適切であったと思われるので同じ時

間を設定した。

スピーキングのタスク：

ABC のテーマのうち2つ (AB, BC, AC) が記入されたプリントが与えられ、どちらか1つテーマを選んで、2分間考えて (語句レベルのメモ可)、2分間でできるだけたくさん自分から話す。どうしても英語が出てこない部分は、日本語の使用を許可。プリントには、図2のような話す内容のヒントが日本語で提示されている。

ライティングのタスク：

スピーキングと同じ3つのテーマが全部記入されたプリントが与えられ (スピーキングと同じ書く内容のヒントが提示されている)、自分がスピーキングで選んだテーマについて、辞書使用なしで20分間自由英作文を書く。スピーキング同様日本語の使用を許可。

▶ 図2：スピーキング・ライティングのテーマ例



このように話す・書く内容のヒントを提示したのは、トピックだけの指定では、何を話したらよいか、何を書いたらよいかわからないため、何も話さない、書かないという問題が起こることが考えられるため、根岸・東京都中学校英語教育研究会 (2007) を参考にした。スピーキングは、英語教員が1人ずつICレコーダーに録音した。なお、スピーキングで長く沈黙が続いた場合 (20秒前後をめぐり)、録音者が質問して発話を引き出すようにした。

4.2 コーパスデータ

上記の324名分のスピーキング (S) とライティング (W) のデータを書き起こし、電子化した。スピーキングには、ah, eh, etto など次の発話を考えているときに発せられるつなぎの音であるフィラーが多数含まれているが、それらは #F# という記号でまとめてコーパス化した。また、スピーキングとライティングの日本語部分は <JP> と </JP> タグで囲んだ。コーパスの概要は表4のとおりであるが、総語数では、フィラーと <JP> </JP> タグで囲まれた日本語部分が1語としてカウントされている。日本語部分は、英語1語に相当するとは限らず、英語で何語かに相当する日本語も多く含まれているので、これを英語1語と同等にカウントし総語数に入れると、総語数としては不正確であるので、下に総語数 (英語のみ) として日本語部分とフィラーを除いた英語のみの産出語数も示した。

■ 表4：コーパスの概要

	中3	高1	高2	高3	計
S&W ファイル数	24	221	38	41	324
S 総語数	853	10,939	3,387	3,499	18,678
平均語数	35.54	49.50	89.13	85.34	57.65
総語数 (英語のみ)	756	9,098	3,100	3,191	16,145
平均語数	31.50	41.17	81.58	77.83	49.83
W 総語数	1,469	11,051	5,041	4,191	21,752
平均語数	61.21	50.00	132.66	102.22	67.14
総語数 (英語のみ)	1,357	9,962	4,989	4,087	20,395
平均語数	56.54	45.08	131.29	99.68	62.95

表5は、各学年のテーマの分布を示している。

■ 表5：テーマの分布

	中3	高1	高2	高3	計
A. ご飯派かパン派か	10	63	11	5	89
B. 訪れたい国	13	107	19	21	160
C. 学校行事	1	51	8	15	75
計	24	221	38	41	324

4.3 分析方法

(1) 英語のみの総語数に関して、スピーキング vs. ライティングの産出平均語数をコーパス全体、英検の熟達度レベル (3級, 準2級, 2級) のそ

それぞれのサブコーパス間で比較した。また、2元配置の分散分析により、モードと英検レベルの間に交互作用があるか調べた。

- (2) コーパス分析ツール (Michael Scott 氏の WordSmith Tools 5.0) を用いて、スピーキングとライティング別に語彙リストを作成し、KeyWords 機能を用いて、語彙リストを相互に統計的に比較し、相対頻度のずれの度合いを対数尤度 (log-likelihood) により検定し、モード間で特徴的な語を抽出した。
- (3) 先行研究でスピーキング vs. ライティングを特徴づける言語特徴として挙げられていた項目を参考にして、表 6 に示した 10 の言語特徴と 12 のサブカテゴリーを選び、それぞれの項目をスピーキングとライティングの全ファイルから抽出し、カイ二乗検定によりモード間で頻度差があるかを調べた。頻度の抽出は、コーパス分析ツール (Laurence Anthony 氏の AntConc の Concordance 機能) を用いて自動検索し、必要に応じて目視判断も行った。検索の際、単語単独で検索可能なものと品詞情報が必要なものがあるが、後者に関しては、形態素解析ツール TreeTagger で品詞タグを付与したデータを用いた。タグ付与の誤りは、一括処理や目視で判断

■ 表 6：分析した言語特徴

A: conjunctions A-1: and A-2: but A-3: so A-4: because	E: N + of
	F: prepositions
	G: subordinate clauses G-1: relative clauses
B: first person references B-1: I	H: passive voice
C: articles C-1: a, an C-2: the	I: present tense I-1: be verbs I-2: lexical verbs
	J: past tense J-1: be verbs J-2: lexical verbs
D: attribute adjectives	

可能なものは、頻度集計の際に数を修正したが、タグ付与の誤りの完全な修正は行っていない。A の接続詞は、節の最初に用いられているもののみをカウントし、語 (句) を結んでいるものは除外した。

- (4) 英検取得者 60 人分のデータを用いて、エラー分析を行った。英検 2 級 (Step 2) 合格者 17 名のうち、1 名のみテーマ A であったので、テーマ A のファイルは除き、残りの 16 ファイル (テーマ B: 11, テーマ C: 5) を用いることにし、それぞれのテーマのライティングの総語数にほぼ等しくなるように、英検 3 級 (Step 3), 準 2 級 (Step 2.5) からファイルをサンプリングした。分析に用いたデータの概要は、表 7 に示したとおりである。

分析項目として、先行研究で中間言語変異が示唆されている項目や (3) の分析結果も踏まえ、頻度情報だけではパフォーマンスの違いを把握できないと思われる項目、さらにデータをざっと観察してエラーが多く見られた項目を総合して、be 動詞、冠詞、名詞の複数形、一般動詞の過去形を選定した。各項目のエラーに下記のようなタグを付与し、コーパス分析ツール (AntConc) を用いて、それぞれのエラーを熟達度レベルごとに各モードで検索し、義務的文脈 (その項目が用いられるべきである文脈) での正用率を算出した。be 動詞は、連結詞 (copula) と助動詞の be を区別した。エラータグは、学習者データを母語話者に修正してもらった修正文を参照しながら手作業で付与した。なお、エラー分析は、エラーの分類基準の恣意性や義務的文脈の特定の難しさなどのため、学習者言語分析としてかなり限界があるという共通認識が第二言語習得研究分野ではあるが、今回使用したデータでは、妥当性に問題が起きるほど判断が分かれそうなケースはほとんどなかったため、本研究では分析方法の 1 つとして行った。

■ 表 7：エラー分析に用いたコーパスデータ

	Step 3			Step 2.5			Step 2		
テーマ	B	C	計	B	C	計	B	C	計
ファイル数	20	8	28	11	5	16	11	5	16
S 総語数	1,173	494	1,667	941	338	1,279	1,108	429	1,537
W 総語数	1,444	583	2,027	1,441	587	2,028	1,443	583	2,026

<エラータグ>

be 動詞: ①<copbe crr=""></copbe>

②<auxbe crr=""></auxbe>

冠詞: ③<at crr=""></at>

複数形: ④<pl crr=""></pl>

過去形: ⑤<pst crr=""></pst>

<エラータグ付与例>

- ① It <copbe crr="was">is</copbe> difficult for us,
so we practiced very hard after school.
- ② I <auxbe crr="was"></auxbe> surprized because ...
- ③ That's <at crr="a"></at> very famous place.
- ④ I want to talk with many <pl crr="foreigners">
forigner</pl>.
- ⑤ I <pst crr="took">take</pst> many pictures.

5 結果

5.1 産出語彙数

表8は、コーパス全体324人のスピーキング vs. ライティングの産出語彙数についての統計量を示したものである。モード間で平均語数（英語のみ）に差があるかt検定を行った結果、統計的に有意差があった（ $p < .01$ ）。同一テーマでスピーキング→ライティングの順で行った本研究のタスクでは、日本人英語学習者は、全体的にライティングの産出語彙数の方が多くなる傾向を示した。

■ 表8：S vs. W の統計量

	平均値	N	標準偏差
S	49.83	324	34.139
W	62.95	324	47.559

表9は、英検の各級・モード別の産出語彙数の基礎統計量を示したもので、図3は、各級のS vs. W

の平均語数をグラフに表したものである。

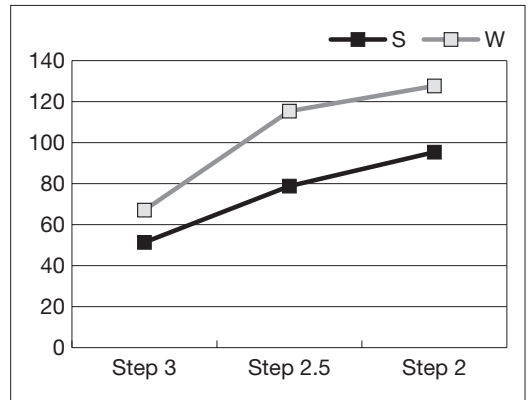
■ 表9：英検レベル別の基礎統計量

	度数	平均値	標準偏差	最小値	最大値
S3	77	51.35	27.140	7	119
S2.5	62	78.79	34.761	30	171
S2	17	95.35	44.633	33	230
W3	77	67.08	29.319	7	145
W2.5	62	115.37	44.533	8	242
W2	17	127.65	32.897	89	220
合計	312	79.96	42.491	7	242

(注) S3, 2.5, 2：スピーキングの英検3級、準2級、2級取得者グループ

W3, 2.5, 2：ライティングの英検3級、準2級、2級取得者グループ

▶ 図3：S vs. W の平均語数



英検の級別を見ると、級が上がるにつれて、両モードとも産出語数が増加傾向にある。熟達度レベルの低い3級では、モード間の語数の差が小さいが、レベルが上がると差が大きくなっている。

総語数に関して、モードと英検レベルの間に交互作用があるか、分散分析を行って調べた結果、表10が示すように交互作用が見られたので（ $p < .05$ ），

■ 表10：分散分析表

ソース	タイプⅢ平方和	自由度	平均平方	F 値	有意確率
mode	40695.582	1	40695.582	34.100	.000
step	136447.916	2	68223.958	57.168	.000
mode * step	7867.538	2	3933.769	3.296	.038
誤差	365181.572	306	1193.404		
総和	2556380.000	312			
修正総和	561499.538	311			

多重比較（Bonferroni による）を行い、6 つのグループ間でどこに差があるかを調べた。多重比較の結果を示すと表11のようになる。

■ 表11：多重比較によるグループ間の差

	S3	S2.5	S2	W3	W2.5	W2
S3		**	**		**	**
S2.5	**				**	**
S2	**			*		
W3			*		**	**
W2.5	**	**		**		
W2	**	**		**		

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

英検の熟達度レベル間では、スピーキングでは3級と準2級、2級の間に有意差が見られるが、準2級と2級の間には有意差は見られない。ライティングでも3級と他の2つのグループは差があるが、準2級と2級の間に差はないというスピーキングと同様の結果を示している。同一学習者集団のモード間では、準2級のみスピーキングとライティングの間で有意差が見られる。異なる学習者集団であるが、スピーキングとライティングの間で総語数に差があることが表11よりわかる。

5.2 特徴語

スピーキングとライティングの各コーパスから作成した使用単語の頻度リストから、ライティングを参照コーパスとして、WordSmith により特徴語を抽出した結果、表12に示す語がモードの違いによる特徴語として抽出された。

■ 表12：モードの違いによる特徴語

Key word	Freq. (S)	Freq. (W)	Keyness
#F#	1254	0	1987.42
I	2188	1970	78.08
FOR	65	164	-30.49
THAT	24	99	-38.36
WAS	113	261	-39.94
IT	121	283	-44.65

(注) Keyness = 対数尤度比の値

#F# はフィラーを表しているが、これはライティングでの使用はないので、スピーキングの特徴を表すものであるが、学習者はつなぎ言葉を多用していることがわかる。もう1つのスピーキングの特徴語は、Iの使用であるが、スピーキングではI like ..., I want to ... のようにIの繰り返しが多いことも頻度差につながっている。ライティングに特徴的な語を見ると、ライティングでは、for, that, it を使い構造が複雑化し、複文的になっていることと、過去形 was の使用が増えていることが特徴である。ライティングでは、for me to ... のような不定詞構文の使用が増えていること、for example を使用して例を挙げ、情報をより詳細に述べていることが特徴として挙げられる。for example は、スピーキングでは24例、ライティングでは49例、for me to ... は、スピーキングでは1例、ライティングでは10例であった。that は、スピーキングでは、指示代名詞としての使用が多いが、ライティングでは、I think (thought) that ... のように接続詞 that の使用が増えており、that 節を用いた複文の使用の増加が目される。

5.3 言語特徴

5.3.1 頻度差

表13は、表6に挙げた10個の言語特徴の粗頻度を調べ、カイ二乗検定を行った結果を示したものである。*は5%水準、**は1%水準で頻度に有意差のあった項目である。

A~Jの10個の言語特徴のうちで、モード間で頻度に関して統計的に有意差が見られなかったのは冠詞のみであったが、サブカテゴリーの定冠詞 the は頻度差があった。調査した言語項目は、a/an, so, be 動詞の現在形、lexical verbs の過去形を除いて、モード間で統計的に有意な頻度差があった。どちらのモードで使用頻度が高いかに関しては、1万語に総語数を標準化して調べてみると、接続詞の but を除いて上で述べた母語話者の先行研究と一致する傾向を示していた。but の使用が書き言葉で多くなっているのは、日本人英語学習者の先行研究（小林, 2008）と一致していた。

■ 表13：言語特徴の頻度

	S：頻度	W：頻度
A: conjunctions**	680	726
A-1: and**	223	188
A-2: but*	72	130
A-3: so	132	146
A-4: because*	253	262
B: first person references**	2430	2283
B-1: I**	2188	1970
C: articles	417	597
C-1: a, an	175	230
C-2: the*	242	367
D: attribute adjectives*	520	750
E: N + of*	93	132
F: prepositions**	625	1136
G: subordinate clauses*	320	483
G-1: relative clauses**	13	55
H: passive voice**	13	41
I: present tense**	1818	2086
I-1: be verbs	596	713
I-2: lexical verbs**	1222	1373
J: past tense**	492	777
J-1: be verbs**	123	295
J-2: lexical verbs	369	482

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

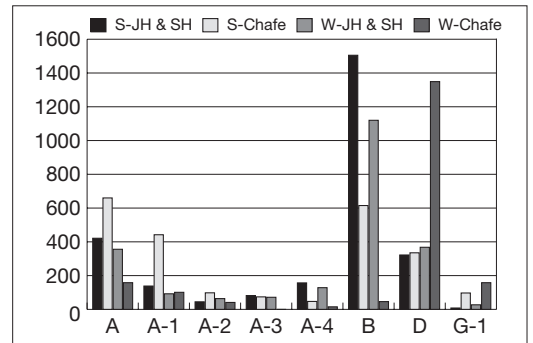
5.3.2 先行研究 (Chafe, 1981) との比較

表13の中で Chafe (1981) で分析されている項目と比較するため、両者の頻度を1万語に標準化した(表14)、視覚化したのが図4である。

■ 表14：日本人中高生 vs. 母語話者の比較 (S vs. W)

	S		W	
	JH & SH	Chafe	JH & SH	Chafe
A: conjunctions	421.18	661	355.97	157
A-1: and	138.12	442	92.18	101
A-2: but	44.60	98	63.74	41
A-3: so	81.76	74	71.59	0
A-4: because	156.70	47	128.46	15
B: first person references	1505.11	615	1119.88	46
D: attribute adjectives	322.08	335	367.74	1349
G-1: relative clauses	8.05	97	26.97	158

▶ 図4：日本人中高生 vs. 母語話者の比較 (S vs. W)



それぞれの項目の左2つが日本人、母語話者のスピーキング、右2つがライティングにおける標準化頻度を表している。Chafeの研究は、日常の食卓での会話と学術論文という言葉のスタイルの違いでは両極端のパフォーマンスを表しているの、同一タスクでの日本人学習者のパフォーマンスとモード間の頻度差が異なる傾向を示していることがわかる。

4つの接続詞を合わせると、話し言葉では母語話者の方が多いが、書き言葉では日本人中高生の方が多い。モード間の頻度差が母語話者は大きく、話し言葉での使用が書き言葉の4倍であるが、日本人学習者の方はモード間の差がそれほどない。母語話者の場合には、書き言葉でbutの使用が話し言葉の約半分になっているのに対し、日本人学習者の場合には、butの使用が書き言葉の方で多くなっている。日本人学習者のsoの使用は、両モードでそれほど頻度差がないが、母語話者ではsoの使用に関して、モードによる違いが顕著である。また、母語話者データでは、日常会話 vs. 学術論文の差を反映して、1人称の代名詞の使用もモード間で頻度差が大きくなっているが、日本人学習者のデータでは、母語話者ほどモード間の頻度差はなく、両モードで非常に使用頻度が高い言語特徴となっている。さらに、属性の形容詞の使用頻度も、母語話者データではモード差が大きい、日本人中高生の場合は、モード間の差は小さい。関係詞節の使用は、日本人の場合は母語話者に比べ、かなり少ないが、それでも書き言葉になると増えている。

以上のように、同じタスクでの日本人学習者のパフォーマンスは、話し言葉と書き言葉の両極端のモード差を示している Chafe の研究と比較すると、モード間の言語特徴の頻度差は小さいが、話し言葉 vs. 書き言葉で見られるモードの違いを全体的には

反映していた。

5.3.3 言語項目の使用状況

調査した項目の使用状況をより詳細に調べてみると、いくつか使用の特徴が見られた。ライティングの方で but が多く使用されている点に関して、データを調べてみると、文頭での But の使用が目立っていた。節の最初に使われている130例の butのうち、80例は文頭での使用であった。これに対し、スピーキングでは、72の使用例のうち20例が文頭での使用であると判断された。なお、スピーキングの書き起こしにおいては、文の切り分けが難しく、butの前で文が終わっているかの判定が難しかったが、その判定は音調を基準として行った。下の例は、ライティングデータの文頭の But を含むコンコーダンスラインの検索結果例である。<s> は文頭、</s> は文末につけているタグである。

▶ 図5：文頭の But を含むコンコーダンスラインの検索結果例

```
peaks English.</s> <s>But the English is different from American Japanese.</s> <s>But, I have to learn American culture, now, Corossam.</s> <s>But there are many heritages in Itarise for dinner.</s> <s>But I sometimes eat <JP>menrui</JP> for it very much.</s> <s>But I like <JP>risotto</JP> and ochazuke it is healthy.</s> <s>But sometimes I don't eat breakfast.</s> <s>But too I love <JP>Yooroppa</JP>.</s> <JP>aji</JP>.</s> <s>But, <JP>takitate</JP> rise like.</s> usually bred.</s> <s>But, dinner is usually rice.</s> <s>My <JP>itoko</JP>.</s> <s>But I will <JP>hanareru</JP> very sad.
```

日本人中高生の because の使用状況を調べてみると、but と同様にライティングで because を文頭で用いている場合が多く見られた。下の例のように、Because を単独の節の最初に文頭で用いている誤用が多く、接続詞の because の使用例の73%は誤用であった。

```
<s>I want to visit America.</s> <s>Because I like English.</s>
<s>I'd like to go to Australia.</s> <s>Because it has beautiful nature.</s>
<s>I like rice.</s> <s>Because it is delicious.</s>
```

これらのそれぞれの文に対応する同じ生徒のスピーキングデータは以下のように書き起こされているため、スピーキングでは誤用にならないが、同じ生徒が S vs. W で同じような文で because を使用し

ていることがわかる。このように、接続詞 because は、話した場合には正しく用いていると思われるでも、書いたデータを見ると使い方を誤っていることがわかった。

```
<s>I want to visit America because I want to study English.</s>
<s>I want to I'd like to Australia because #F# it has beautiful nature.</s>
<s>I like I like rice because it is delicious.</s>
```

1 人称の代名詞 I の使用状況を見てみると、5.2 特徴語の節でも述べたように、スピーキングで I の繰り返しが多いのが特徴である。次の例は、I の繰り返しの検索結果例と同一学習者の S vs. W における I の使用例である。

▶ 図6：I の繰り返しの検索結果例

```
blue.</s> <s>#F# I I feel it is beautiful.</s> <s>...</s>
/s> <s>...</s> I I want to <./s> I go to shopping because shoes.</s> <s>I I want to stay France in future.</s> <s>
kt> <s>...</s> I I went I went to Okinawa.</s> <s>...</s>
kan</JP>.</s> <s>I I for fishing fishing.</s> <s>Q</Q></s>
yo Kyoto.</s> <s>I I love I love Kyoto's Koyoto's maiko.</s>
's maiko.</s> <s>I I want I want to I want to meet meet ma
enty-six.</s> <s>I I will I will speak I don't speak <./s>
/s> <s>...</s> I I play with American friends.</s> <s>...
.</s> <s>...</s> I I like best <./s> curry rice <./s>
```

<スピーキング>

```
I want to go to America #F# because I I want to I
want to speak English very well.
```

<ライティング>

```
I want to go to America.
Because I want to learn American culture, and to
speak English very well.
```

ライティングでは、I に続ける内容を考えてから文字化できるが、時間的余裕のないスピーキングではまず I と言って、I を繰り返しながら、続ける発話の内容を考えている。コンコーダンスラインの使用例では、I want to や I will という句の繰り返しも見られる。また、同一学習者の例には、上で述べた because の誤用も含まれている。

関係詞の使用は、スピーキングでは極めて限られているが、ライティングでは使用例が増えており、後置修飾による構文の複雑化が見られた。次の例

は、同じ生徒のスピーキング vs. ライティングでの関係代名詞の使用例（下線部）である。スピーキングでこのように関係代名詞を使用している例は少ないが、この生徒はライティングではさらに多く関係代名詞を使った文を書いていた。

<スピーキング>

- ① But I want to go to England, because England England there is a lot of place which which which I want to visit in England.
- ② And another country which I want to visit is #F# when when I get college student I want to I want to go to for a country which is which people cannot study because #F# the country the country's educational system is not not good.

<ライティング>

- ① The countries which I want to visit are England, America, Australia and so on.
- ② The places which I want to see in England are Oxford, <JP>ruuburu</JP> museum, and so on.
- ③ ... when I become a university student, I'd like to go to countries which the educational system is bad.
- ④ But I don't know about the program which helps poor children with studying.
- ⑤ But country which I want to go to as a sightseeing is Australia.
- ⑥ These are countries which I want to visit.

上の例には、関係詞の誤用もあるが、同一学習者の産出例として興味深いものである。スピーキングの①②の発話内容は、ライティングでは②③のように表現されている。スピーキングでは、考えながら関係代名詞を繰り返し、節をつなげていこうと苦勞しながら話していることがうかがえるが、ライティングになると考えをまとめて整理して、関係代名詞を活用しながら書いていることがわかる。

5.4 エラー分析

表15は、英検の熟達度レベルごとに、連結詞（copula）の be、助動詞（aux）の be、冠詞、名詞の複数形、一般動詞の過去形について、これらの項目が使用されなければならない義務的文脈におけるエラーの数および義務的文脈に占めるエラーの割合を示したものである。右側の（ ）内の数字は義務的文脈の数である。

義務的文脈数が少なかった助動詞としての be 動詞を除けば、その他の項目では、各レベルにおいて、ライティングの方でその項目が使用されるべき文脈数が多い傾向にある。コーパスサイズがほぼ同じであるライティングの方で、熟達度レベル間でエラーの状況を見てみると、連結詞の be、複数形、過去形のエラー数はレベルが上がると次第に減少している。表15の義務的文脈数には、その項目が必要な箇所に抜けている脱落のケースが含まれ、不必要な箇所に余分に使用されている余剰のケースは除かれているため、データにおける項目の生起頻度とは異なる。特に、冠詞に関しては、必要な箇所に冠詞を付け忘れているケース（脱落タイプ）が数多く見られたため、義務的文脈の数は実際の冠詞の生起頻度より多くなっている。表16は、冠詞の使用頻度およびエラーの状況をモード別、熟達度レベル別に示したものである。（ ）内は、エラー数に占めるそれぞれのタイプのエラーの割合である。

冠詞のエラータイプは、表16に示したように、脱落タイプが占める割合が最も高いが、英検2級のデータでは、他の級と比較すると、両モードで、脱落タイプの割合が低くなり、置換タイプ（冠詞が必要な箇所に、誤った冠詞、あるいはその他の決定詞が付けられているケース）のエラーの割合が増えている。モード間で比較すると、英検3級のデータでは、両モードで共に余剰タイプの割合が2番目に高いが、英検準2級および2級のデータでは、S vs. W

■ 表15：エラー数

	Step 3		Step 2.5		Step 2	
	S	W	S	W	S	W
copula be	13(25.0%) (52)	20(23.0%) (87)	12(29.3%) (41)	13(16.7%) (78)	10(25.0%) (40)	7(12.5%) (56)
aux be	2(13.3%) (15)	2(11.8%) (17)	1(8.3%) (12)	3(25.0%) (12)	6(35.3%) (17)	3(18.8%) (16)
articles	32(47.1%) (68)	43(44.8%) (96)	19(44.2%) (43)	52(57.8%) (90)	36(46.8%) (77)	29(34.5%) (84)
plural	20(58.8%) (34)	31(41.9%) (74)	12(48.0%) (25)	30(41.1%) (73)	20(46.5%) (43)	22(28.2%) (78)
past tense	15(34.9%) (43)	11(19.3%) (57)	7(19.4%) (36)	10(16.1%) (62)	5(15.6%) (32)	3(5.4%) (56)

■ 表16：冠詞のエラー状況

	Step 3		Step 2.5		Step 2	
	S	W	S	W	S	W
生起頻度	44	66	28	43	53	66
義務的文脈	68	96	43	90	77	84
脱落タイプ	29 (78.38%)	40 (75.47%)	19 (82.61%)	48 (90.57%)	32 (72.73%)	22 (66.67%)
余剰タイプ	5 (13.51%)	10 (18.87%)	4 (17.39%)	1 (1.89%)	8 (18.19%)	4 (12.12%)
置換タイプ	3 (8.11%)	3 (5.66%)	0 (0%)	4 (7.55%)	4 (9.09%)	7 (21.21%)
エラー数計	37	53	23	53	44	33

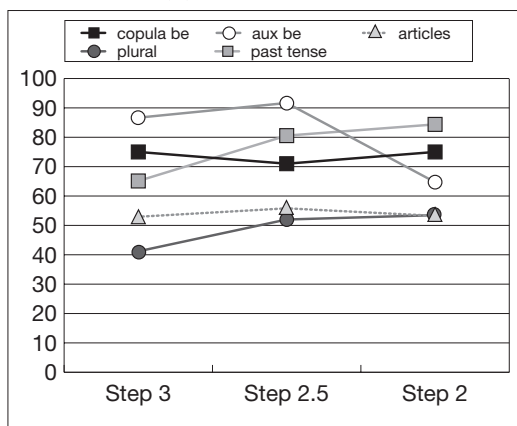
■ 表17：正用率

	Step 3		Step 2.5		Step 2	
	S	W	S	W	S	W
copula be	75.00%	77.01%	70.73%	83.33%	75.00%	87.50%
aux be	86.67	88.24	91.67	75.00	64.71	81.25
articles	52.94	55.21	55.81	42.22	53.25	65.48
plural	41.18	58.11	52.00	58.90	53.49	71.79
past tense	65.12	80.70	80.56	83.87	84.38	94.64

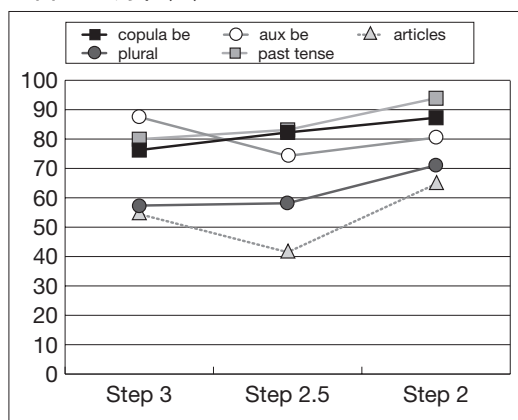
で余剰タイプと置換タイプの割合が逆転し、ライティングの方で置換タイプの割合が2番目に高くなっている。特に2級のライティングで置換タイプの割合が高くなっており、脱落タイプの割合が低くなっている。

表17は、表15に基づき各項目の正用率を算出した結果を示したものであり、モード別にグラフに表したものが図7、図8である。

▶ 図7：正用率 (S)



▶ 図8：正用率 (W)



正用率をモード間で比較すると、連結詞の be、名詞の複数形、一般動詞の過去形では、各熟達度レベルでライティングの方が正用率が高い。連結詞の be は、英検3級レベルでは、モード間で正用率の差が小さいが、準2級および2級レベルでは、モード間で差が大きくなっている。複数形と過去形は、3級および2級レベルでモード間の正用率の差が大きい、準2級レベルでは、他のレベルと比べると差が小さい。これら2つは、熟達度レベルが上がるにつれ、正用率が上昇している項目であるが、モードの違いで異なる上昇傾向を示している。スピーキングでは、両項目とも3級から準2級になると10%

以上正用率が高くなっているが、準2級と2級ではあまり差がない。一方、ライティングでは、3級と準2級では正用率の差が小さいが、2級になると準2級よりも10%以上正用率が高くなっている。

冠詞は、各レベルでモード間の正用率の差に異なる傾向が見られる。レベルが上がっても、正用率はスピーキングでは50%台と低いままである。一方、ライティングでは、2級レベルで正用率が最も高くなっているが、約65%でレベルが上がっても依然としてエラーが多い項目である。助動詞の be は、上で述べたように、義務的文脈数が少ないため、このデータからはその傾向を述べることは難しいであろう。2級レベルでは、すべての項目において、ライティングの方で10%以上正用率が高くなっていることが注目される。

6 考察

本研究では、スピーキング→ライティングの順で中高生の同一学習者集団からデータを収集したが、産出語彙数の平均は、コーパス全体では、ライティングの方が10語余り多く、統計的にも有意な差があった。これは、スピーキング→ライティングの順で、スピーキングの準備時間2分間＋スピーキング2分間、ライティング20分間という条件の下でのデータ収集の結果であり、planning timeの違いや practice effect が総語数に影響している可能性があるが、このような条件の下では、総語数がスピーキング vs. ライティングパフォーマンスの差を示す指標になりうると言える。

また、モード間の産出語彙数の差には、学習者集団の英語のレベルや日頃の授業での指導、活動が影響を与えているように思われる。日常的に英語を(即興で)話す活動、書く活動をしているかどうか、まとまった量の英文を話す、書く練習をしているのかなどが、モード間で生じる語数の差に関係しているかもしれない。例えば、まとまった量の英語を即興で話す活動も書く活動も日常的にしていない比較的英語のレベルが高い学習者集団とレベルが低い学習者集団では、前者はモード間の産出語数の差が大きくなり、後者ではそれほど差が出ない傾向が見られた。同じように日常的にまとまった英文を書く練習はしていなくても、英語力の差で産出量に差

が出たと考えられる。

一方、高2の学習者集団のように、話したり書いたりする練習を日常的に行っていて、英語のレベルも高い集団は、モード間の差が約50語と大きくなっていたが、このような集団では、即興で話すこともある程度でき、さらに planning time がある中で書いた場合にはかなり多くの量を産出することができると推測される。

特徴語の分析で抽出された語を見ると、先行研究で英語の話し言葉 vs. 書き言葉の差異を示している1人称の代名詞 I や過去形の was が抽出されていること、for, that, it が書き言葉での構造の複雑化を示唆していることは、特徴語がスピーキング vs. ライティングパフォーマンス比較分析のための指標として機能する可能性があることを示している。ただし、本調査で代名詞 I が話し言葉で多かったのは、I の繰り返しが多く見られたことから、単にモード差によるものではなく、流暢さの問題も関係していると思われる。

10個の言語特徴をスピーキング vs. ライティングパフォーマンスで比較分析した結果、母語話者が一般に示す産出モードの違いによる特徴を示していたものについては、日本人中高生の学習者集団のパフォーマンス比較にも有効な指標であると言えるだろう。

ただし、話し言葉と書き言葉の両極端のスタイルを比較した Chafe (1981) の研究との比較で、各項目のモード間の頻度差の程度において、母語話者が示す差とは異なる傾向を示していたことからわかるように、収集したデータのスタイルの違いにより、モード間で現れる頻度差は推移することが予測される。Chafe との比較で、4つの接続詞 (and, but, so, because) の使用頻度が、それぞれのモードで母語話者と異なることが示されたが(4つを合わせると、話し言葉では母語話者の方が多いが、書き言葉では日本人中高生の方が多い)、これは、日本人と英語母語話者の接続詞の使い方に関する基本的な違いによるものであると考えられる。つまり、日本人が日本語で文章を書くと、英語母語話者が英文を書く場合よりも接続詞の使用が多い傾向にあり、書き言葉におけるこれらの接続詞の多用は、母語の日本語からくる特性であると思われる。また、日本人英語学習者は、使える接続詞のレパートリーが母語話者と異なり、限られているので、書き言葉でもこ

これらの4つを多用する傾向があると言えるだろう。

大学生の話し言葉 vs. 書き言葉を調査した朝尾(2008)は、日本人英語学習者の書き言葉は話し言葉的であるという特徴を指摘し、学習者の言葉は、話し言葉、書き言葉が未分化の状態であると述べている。本研究では、未分化であるとはいえないが、接続詞や1人称代名詞のIなどの使用頻度差を見ると、書き言葉が話し言葉の特徴を有していると言えるであろう。とりわけ、becauseの単独の節の文頭での使用に見られるように、話し言葉と書き言葉で異なる文法があまり意識されていないことがうかがえる。

スピーキングでは、関係代名詞を繰り返し、苦労しながら使っていたが、ライティングでは、うまく考えをまとめて関係代名詞を駆使していた使用例は、planning timeやmonitoringが影響して生じた変異だと考えられ、話し言葉では使用が難しい項目や複雑な構造も時間が十分ありmonitoringが働く書き言葉になると、うまく使えると予想される。

母語話者では通常産出モードの違いにより差異が見られる項目で、差異があまり見られなかったものは、日本人英語学習者あるいは本研究に参加した学習者集団独自の特徴であるかもしれない。冠詞全体ではモード間で統計的に有意な差が見られなかったが、冠詞はエラー分析で示したように、エラーが多かった項目であり、頻度情報だけではパフォーマンスの違いを判断することは難しく、エラーの状況や正用率も含めて比較して見る必要があるように思われる。本研究の分析は、英検取得者のみのデータで行い、熟達度レベル間で、冠詞の正用率は一定した傾向を示していなかったが、2級レベルでは、モード間の正用率に差が見られたことから、ある一定のレベル以上になると、書き言葉の方で正用率が高くなる安定した傾向を示すのかもしれない。さらなる検証が必要である。

同様に、頻度情報では有意な差を示していなかったbe動詞(現在形)や一般動詞の過去形なども、正用率がモード差を示す指標になる可能性があるであろう。2級のデータでは、エラー分析を行ったすべての項目で、安定して10%以上書き言葉の方で正用率が高くなっていたが、これは、planning time, monitoring, practice effectなどが影響した結果であると考えられる。英検2級レベルの学習者集団では、分析した5項目については、正用率がパフォー

マンスの違いを表す指標になりうることが示唆されている。他の言語項目についても調査してみると、より多くの指標が明らかになると思われる。

7 まとめと今後の課題

本研究では、日本人中高生324人からスピーキングとライティング両方のデータを同一テーマで収集してコーパス化し、スピーキング vs. ライティングパフォーマンスを比較分析するのに有効な指標を調べた。話す→書くという順でタスクを実施して得られた産出データを用いて、使用語彙、10の言語特徴(12のサブカテゴリー)、5つの項目のエラーおよび正用率について、話し言葉 vs. 書き言葉で比較分析を行った。

参加者全員の平均語数では、書き言葉の方が多く、モード間に有意な差が見られた。英検取得者のサブコーパスを調べると、産出語数に関してモードと熟達度レベル(英検の級)の間で交互作用があった。それぞれのモードで使用されている語彙リストを比較して抽出された特徴語は、モードの違いによる特徴を反映していた。10の言語特徴の比較分析結果では、冠詞を除いて9つの言語特徴で使用頻度においてモード差が見られた。エラーを調べた項目についての正用率については、すべての熟達度レベルではないが、モードの違いで変異が見られた。

以上の分析結果から、中学生・高校生の日本人英語学習者のスピーキング vs. ライティングパフォーマンスの比較分析のための有効な指標として、総語数、特徴語、接続詞(and, but, so, because)、1人称代名詞、冠詞(頻度+正用率)、属性の形容詞、N+of、前置詞、従属節、受動態、現在形、過去形(頻度+一般動詞の過去形の正用率)、連結詞beの正用率、名詞の複数形の正用率を挙げることができる。

日本人英語学習者の同一学習者集団の産出モードの違いによる中間言語変異の研究には、これまで大規模なデータを用いた研究はなく、特に中高生の実態はまだほとんど解明されていない。本研究は、中高生の中間言語の実態解明を試みた研究として有益であり、収集したデータの価値は高いものであろう。ただ、本研究で用いたデータは、学年、英検レベルにより収集数に偏りがあるので、今後ある程度バランスがとれるようにデータの補充を検討することも

必要であろう。今回は、データ収集とコーパス化にかなりの時間を要したこともあり、コーパス全体と英検取得者のサブコーパスを用いて、先行研究でモード差が示されている項目に焦点を当てて調べたため、多項目は扱っていない。今後より多くの項目を学年別、英検の熟達度レベル別、個人間および個人内などで分析してみることににより、新たに指標となる項目が見いだされる可能性がある。エラー分析については、限定した項目しかまだ調査していないので、今後より多くの項目を調査することにより、指標となる項目が明らかになることが期待される。

日本人中高生のスピーキング vs. ライティングパフォーマンスの比較から明らかになった特徴を今後どのように英語教育で応用することができるかを考えなければならない。新学習指導要領が重視している技能間の有機的なつながりを考慮に入れた指導を行う際、ある話題について話す→書くという組み合わせの指導が1つの有効な方法であると考えられるので、そのような指導を行うにあたって、本研究の調査結果を生かすことができるであろう。例えば、今回の調査で明らかになった中高生の産出語彙数の

実態を参考にして、各学年におけるスピーキングとライティングの目標産出語数を定め、各発達段階で焦点を当てる言語項目を選定し、モードによる使用頻度、使用状況の特徴、正確度などの差異を考慮に入れて指導を行うことが可能になるであろう。今後、さらに調査を進め、指導法や教材化、スピーキング vs. ライティングパフォーマンスの測定に調査結果を応用していくことが期待される。

謝 辞

本研究の機会を与えてくださいました(財)日本英語検定協会の皆様、選考委員の先生方、特にまとめの段階で貴重なコメントをいただきました大友賢二先生に厚くお礼申しあげます。また、データ分析やまとめに関して有益なアドバイスや示唆を与えてくださいました東京外国語大学大学院の投野由紀夫先生と吉富朝子先生に心より感謝申しあげます。最後に、本研究は、中高の英語科の多くの先生方と生徒の皆様のご協力なくしては実現できませんでした。心よりお礼申しあげます。本当にありがとうございました。

参考文献 (*は引用文献)

- * Abe, M. (2003). A corpus-based contrastive analysis of spoken and written learner corpora: The case of Japanese-speaking learners of English. In D. Archer et al. (Eds.), *Proceedings of the Corpus Linguistics 2003 Conference. Technical Papers 16* (pp.1-9). Lancaster University: University Centre for Computer Corpus Research on Language.
- * Abe, M. (2007). Grammatical errors across proficiency levels in L2 spoken and written English. *The Economic Journal of Takasaki City University of Economics*, 49(3, 4), 117-129.
- * 朝尾幸次郎. (2008). 「中間言語は学習者コーパスにどのように現れるか」. 中村純作・堀田秀吾 (編). 『コーパスと英語教育の接点』 (pp.91-104). 東京: 松柏社.
- * Biber, D. (1986). Spoken and written textual dimensions in English: Revolving the contradictory findings. *Language*, 62, 384-414.
- * Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.
- * Chafe, W.L. (1981). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp.35-53). Norwood, NJ: Ablex.
- Corder, S.P. (1967). The significance of learner's errors. *IRAL*, 5(4), 161-170.
- Corder, S.P. (1981). *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- * Ellis, R. (1987). Interlanguage variability in narrative discourse: Style-shifting in the use of the past tense. *Studies in Second Language Acquisition*, 9, 1-20.
- * Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- * Haynes, L. A. (1992). The development of speaking/writing variability in narratives of non-native English speakers. *Issues in Applied Linguistics*, 3(1), 43-67.
- * 和泉絵美・内元清貴・井佐原均 (編著). (2004). 『日本人1200人の英語スピーキングコーパス』. 東京: アルク.
- Kimura, M. (2004). *Describing Japanese EFL Learners' Performance and the Process of its Development: Utilizing a Large-scale Learner Corpus and Data-driven Techniques*. Unpublished PhD thesis. Tokyo Gakuhei University.
- * 小林雄一郎. (2008). 「高頻度語を用いた学習者コーパスの分類」. 『学習者コーパスの解析に基づく客観的

- 作文評価指標の検討』(統計数理研究所共同研究リポート215). 東京: 統計数理研究所, 69-82.
- * 小林雄一郎.(2009).「コーパス言語学における判別分析の応用」.『コーパス言語研究における量的データ処理のための統計手法の概観』(統計数理研究所共同研究リポート232). 東京: 統計数理研究所, 39-52.
- * 小林雄一郎・山田洋文.(2008).「日本人英語学習者コーパスに見る Metadiscourse Markers の使用傾向」.『英語コーパス研究』第15号, 161-173.
- 小池生夫(編集主幹).(2003).『応用言語学事典』. 東京: 研究社.
- * 小泉利恵.(2007).「スピーキング・テストにおける語彙的複雑さの指標とその特徴」. 第46回大学英語教育学会全国大会「JACET 英語語彙研究会企画シンポジウム 語彙の測定とその問題点: さまざまな測定方法によって引き出される語彙の側面」. 発表資料.
- Krashen, S. (2003). *Explorations in Language Acquisition and Use*. Portsmouth, NH: Heinemann.
- * Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, 9(4), 409-419.
- McEnery, T., Xiao, R. & Tono, Y. (2006). *Corpus-Based Language Studies: An Advanced Resource Book*. Abingdon: Routledge.
- * 三浦愛香.(2008).「会話(NICT JLE) vs. 作文(JEFL) コーパスの比較と分析: 英語学習段階と名詞の内部構造発達」.『英語コーパス研究』第15号, 135-148.
- * 水本篤.(2008).「自由英作文における語彙の統計指標と評定者の総合評価の関係」.『学習者コーパスの解析に基づく客観的作文評価指標の検討』(統計数理研究所共同研究リポート215). 東京: 統計数理研究所, 15-28.
- 文部科学省.(2008).『中学校学習指導要領解説 外国語編』. http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/chukaisetsu/index.htm
- 文部科学省.(2009).『高等学校学習指導要領解説 外国語編 英語編』. http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/1282000.htm
- Nation, I.S.P. (2008). *Teaching Vocabulary: Strategies and Techniques*. Boston, MA: Heinle Cengage Learning.
- * 根岸雅史・東京都中学校英語教育研究会(編著).(2007).『コミュニケーション・テストングへの挑戦』. 東京: 三省堂.
- * 野田哲雄(代表).(2008).『日本人英語学習者のスピーキング能力の実態調査ならびに評価指標開発—中学生・高校生を対象として』. 平成17年度—平成19年度科学研究費補助金(萌芽研究)研究成果報告書, 課題番号 17652059.
- * 野村真理子.(2009).「日本人英語学習者の産出モードの違いによる言語特徴の分析—中学・高校段階の学習者に焦点をあてて」.『四国英語教育学会紀要』第29号, 15-24.
- * Ochs, E. (1979). Planned and unplanned discourse. In T. Givon (Ed.), *Discourse and Syntax* (pp.51-80). New York: Academic Press.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- * Selinker, L. (1972). Interlanguage. *IRAL*, 10(3), 209-231.
- * 鈴木理恵・投野由紀夫.(2008).「英語学習段階と基本動詞の意味・構造発達: 動詞 *get* に着目して」.『英語コーパス研究』第15号, 149-160.
- * Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language*, 58(1), 1-21.
- * Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning*, 32(1), 69-84.
- * Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, 4(2), 142-164.
- * Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning*, 35(3), 373-403.
- Tono, Y. (2002). *The Role of Learner Corpora in SLA Research and Foreign Language Teaching: The Multiple Comparison Approach*. Unpublished PhD thesis. Lancaster University.
- * 投野由紀夫(編著).(2007).『日本人中高生一万人の英語コーパス JEFL Corpus—中高生が書く英文の実態とその分析』. 東京: 小学館.
- * 投野由紀夫.(2008).「NICT JLE vs. JEFL: n-gram を用いた語彙・品詞使用の発達」.『英語コーパス研究』第15号, 119-133.
- * Vann, R.J. (1979). Oral and written syntactic relationships in second language learning. In C. Yorio, K. Perkins & J. Schachter (Eds.), *On TESOL '79* (pp.322-329). Washington, D.C.: TESOL.
- * Young, R. (1988). Variation and the interlanguage hypothesis. *Studies in Second Language Acquisition*, 10, 281-302.
- * (財) 日本英語検定教会.(2006).『英検 Can-do リスト』.