

静岡産業大学 情報学部 研究紀要  
Journal of Shizuoka Sangyo University  
第24号 (2022) 別刷

ある英語学習者における英語ライティングの変化  
—検定資格試験学習を通して—

Longitudinal Changes in an English Language Learner's Writing Performance:  
In the Course of Studying for the English Proficiency Test

季 駿 法月 健

Shun KI Ken NORIZUKI

静岡産業大学 情報学部



## ある英語学習者における英語ライティングの変化 —検定資格試験学習を通して—

Longitudinal Changes in an English Language Learner's Writing Performance:  
In the Course of Studying for the English Proficiency Test

季 駿 法月 健  
Shun KI Ken NORIZUKI

(令和3年10月5日受理)

本研究の目的は、個人学習者の英文ライティングを、数量的及び質的に分析することで、学習者のライティングがどのように変化したのかを考察することにある。近年ライティング分析等への活用が報告されているオンライン分析ツール、TAALES、TAACO、TAASSCとMicrosoft Wordのツールから得られた指標に対して、一元分散分析を行ったところ、文章の読みやすさ（リーダビリティ）の指標であるFlesch Reading Ease (FRE)や、文章の統語的な複雑さの指標であるMLC（節内の平均語数）等から有意差が見つかった。これらの結果を、実際の学習者のライティングと照らし合わせて、質的な分析を行ったところ、文章に使用されている語がより難解になり、文構造がより複雑になる傾向がわかった。また、名詞句を修飾する表現や言い換え（パラフレーズ）の表現がより多くなったことも明らかになった。さらに、FREとMLCをライティング得点の代替となる指標（目標変数）と仮定して、上記のツールから得られた様々な指標の組み合わせを説明変数として重回帰分析を行った結果、3つの有効な予測式が得られた。

**キーワード：**個人学習者、英文ライティング、オンライン分析ツール、数量的分析、質的分析

### 1. はじめに

現代人の生活において、文章を書くこと（ライティング）は、学業、職業、社会的な交流の様々な活動において、重要な役割を担っている。外国語のライティングにも同じことがあてはまるが、分野を絞って考えても、その能力を構成する要素は多岐にわたり、客観的かつ多面的に評価を行い、適切な指導や学習に結び付けていくことは決して容易とは言えない。

近年、英語の文章や発話を解析する無料のオンライン分析ツールが普及し、使用語彙の多様性や洗練性、統語的洗練性や複雑性、論理的な結束性等、ライティングやスピーキング、リーディングの様々な目的の分析に活用されている（例、Crossley, Kyle & McNamara, 2016a; 2016b; Kyle, 2016; Kyle & Crossley, 2015; Hamada, 2015; 佐々木, 2020; 小室, 2020）。しかしながら、これらの研究は、特定の学習者集団の全体的な特徴をとらえることを目的とし、個別学習者に還元できるような研究成果は限られていることが多い。すべての学習者や異なる活動テーマに対して、万能な効果を発揮するような指標はなく、膨大

なデータが個別の学習者に十分な解説がないまま提供されても、効果的に活用されることは期待できないだろう。

Meara (1995) は、語彙習得の分野における単一被験者の学習を追究する研究の重要性を唱え、自らの語彙習得について研究した 19 世紀の研究者の取り組みが見直されるべきであると述べているが、オンライン分析ツールの支援が得られるようになった現在、個人学習者の多面的な英語能力の推移に焦点を当てることの意義は、もっと評価されるべきかもしれない。

本論文の第一執筆者は、大学の学部生時代に、ゼミ研究の活動の一貫として実用英語技能検定（以下、英検）の準 1 級取得に励み、計 4 回の受験の末、合格することができた。当初は、検定試験に合格することが目標であったが、徐々に練習を積み重ねていくうちに、単なる資格の取得だけではなく、その先にある英語を使って自分の表現したい事柄を受け手にどのように伝えるべきかを試行錯誤する過程に重点を置くようになった。これは、英検の二次面接であるスピーキングテストへの練習がきっかけとなったことだが、同じ表現や言い回しが文章にも活用できることから、ライティングへと波及していった。実際、ライティングテストの練習では、同じトピックに対して期間をおきながら何度も書いていたり、トピックに関する立場を賛成と反対の両方で論じたりと、文章の表現方法を模索したり、多角的にトピックを考えたりすることが最終的に増えていった。

また、学部卒業前には、少人数の学部生が執筆した英文ライティングのスペルエラーの特徴について、数量的な分析を行い、学会で発表を行った（季, 2021）が、一学習者として、自らの英検ライティングの練習から得た知見が大いに参考になった。とりわけ、実験に参加してもらった協力者に対して、ライティングのフィードバックを行ったことが、自らのライティングを振り返るきっかけとなった。フィードバックではスペルエラーの他に、より自然な表現の示唆や論理的構成にするための提案も含めた。他の英語学習者のライティングを読み、添削する機会を得たことで、改めてこれまで行ってきた自身のライティング活動を振り返り、それを客観的に分析することで、どのように表現方法や構成が変わっていったのかを考察することに興味をもったことが本研究の動機の一つになっている。

本研究では、一人の学習者が英検準 1 級取得のために約 1 年間にわたって執筆した合計 41 編の英文ライティングを、数量的及び質的に分析する。具体的には、「学習者」のライティングを語彙や文法、文章の結束性など様々な指標を提供するツールを用いて数量的に分析し、そこで有意な結果となった指標を実際のライティングと照らし合わせることで、「学習者」のライティングにどのような変化が生じたのかを質的に考察する。

## 2. 先行研究

近年、コンピュータプログラムを使って、ライティングの自動採点が行われることが多くなっているが、有料で、教員や学習者が自由に使用することができない場合が多い。その一方で、無料で活用できるウェブ上の文章解析のツールや、自由にダウンロードして使用できるプログラムが、急速に普及しており、英語学習の様々な側面について、ツールを介した分析が行われるようになってきている。

Hamada (2015) は、Coh-Metrix というツールを使って英検の 1～3 級の読解問題を

分析した結果、語彙や文法指標のような表層的な言語学的変数が、より高次なものよりも正確に、英検のテスト級の識別を予測できたことを報告している。Coh-Metrix と Text Inspector を英検 1～3 級のライティング模範解答の分析に応用した佐々木 (2020) は、受験級が高くなるほど、トピックが難しくなり、幅広い単語や時制・相を用いたライティングが求められるとしている。

一方、Salsbury, Crossley & McNamara (2011) は、6 人の英語学習者の 1 年間のスピーキングコーパスの比較を Coh-Metrix の語彙指標である具体性 (concreteness)、心像性 (imageability)、親密性 (familiarity)、有意味性 (meaningfulness) について行ったところ、親密性を除いて、月日が経つにつれて、数値が有意に下がったことを確認した。これに対して、Crossley & Skalicky (2019) は、同じ指標を TAALES と呼ばれるツールを使って、TESOL (外国語としての英語教育) 専攻の学部生及び大学院生計 50 名 (半分は母語話者) と大学の語学プログラムで学ぶ学生 50 名の会話を分析したところ、時期が経つにつれて、Salsbury, et al. (2011) と同じ指標に有意な減少があり、英語力の高い学生ほどそれらの数値が低くなった。

つまり、習熟度が高くなるにつれて、抽象度の高い単語、視覚的にイメージしにくい単語、他の単語との意味的な関係性がとらえにくい単語を多用するものの、スピーキングで使用する語彙の親密性にまでは影響が出なかったようである。

Coh-Metrix は、一回の使用で単一の文章しか分析できないが、語彙の洗練性測定の TAALES、文章の結束性測定の TAACO、統語的洗練性・複雑性測定の TAASSC は、複数の文章をフォルダーに入れて、バッチ処理で、同時に解析することができる。

Coh-Metrix に比べてより分野別に特化された TAALES や TAACO、TAASSC 等は、簡易に大量のデータをバッチ処理できる特性のためか、分析に併用されることが増えてきているようだ。併用することで、大量の指標を入手することが容易にできるようになったが、800 近い指標のすべてについて、様々な検証を行うことは現実的とは言えず、活用する教員や学習者に、事後の分析や判断を委ねることもできない。

個人学習者の日常的なライティング活動を分析する本研究においては、一般的なグループ研究と異なり、各エッセイに付与されたスコアやその他の対比できるテストスコアもなかったが、関連研究の経時的な比較の中で、大きな変化を見せた指標 (Crossley & Skalicky, 2019; Salsbury, et al., 2011 等) や重回帰分析等を通じてスコアや習熟度を表す級の予測に効果が見られた指標 (Crossley, Kyle & McNamara, 2016b; Jung, Crossley & McNamara, 2019; 小室, 2020)、本研究の「学習者」の過去の作品に対する懐古的、質的な分析から得られたフィードバックを参考にして、分析を行う。まず、経時的な変化等が顕著な指標を探り、その中で特に安定して変化しているものをテストスコアに代わる目標変数と仮定して、それと何らかの関連性を持つことが予想される指標 (説明変数) との間で、重回帰分析を行う。最後に、数量的分析から得られた結果を基に、質的な分析を行う。

### 3. 調査

#### 3.1 目的

「学習者」が英検準 1 級対策で、2018 年夏から 2019 年秋にかけて、作成したライティ

ングを分析の対象とした。具体的にはTAALES、TAACO、TAASSCの指標の中で、ライティングの経時的な比較を行った結果、数値に大きな差が検出されたものを探り、文章の複雑さを示す指標をライティングのスコア（目標変数）と仮定した場合に、スコアを予測する最も効果的な説明変数の組み合わせを算出し、それらの数値とライティングの関係を質的に分析する。研究課題（Research Questions: RQ）は以下の通りである。

- RQ1: 個人学習者のライティングの TAALES、TAACO、TAASSC の指標の中で、学習期間を経て、顕著に変化したものは何か。
- RQ2: 個人学習者のライティングの複雑さの数値を、ライティングのスコアと仮定した場合に、そのスコアを効果的に予測する指標の TAALES、TAACO、TAASSC の指標の組み合わせは何か。
- RQ3: 個人学習者のライティングの質的な変化に、TAALES、TAACO、TAASSC の指標で反映されているものはあるか。

## 3.2 方法

### 3.2.1 ライティング

2018年夏季から2019年秋季にかけ、英検準1級取得を目指して「学習者」が書いた41本のライティング作品（エッセイ）を電子ファイル化した。使用したライティング・トピックの内訳は、英検の過去問題21本、英検のウェブサイトで紹介された予想問題7本、学習者が独自に考えたオリジナル問題13本であった。

41本の作品のうち26本は、同じトピックで2回以上書かれていて、そのうちの21本が3回以上、そのうちの4本は4回書かれたものである。繰り返し書かれた26本の中には、間隔を空けて書かれたものも、短期間で連続的に書かれたものも含まれる。

英検準1級のライティングのトピックは、大きく賛成／反対意見型（Agree or Disagree: A should do/be ...）、予測型（Do you think that A will do/be ...）に分かれ、いずれも4つのキーワード（フレーズ）から2つを選んで、120～150語で、導入、本体、結論の構成で、作成することが求められている。

### 3.2.2 分析ツール

41本のエッセイを、テキストファイルの形式で、バッチ処理のための分析用のフォルダーに入れ、語彙洗練性分析のTAALES (the Tool for the Automatic Analysis of Lexical Sophistication) (Kyle & Crossley, 2015)、テキスト結束性分析のTAACO (the Tool for the Automatic Analysis of COhesion) (Crossley, Kyle & McNamara, 2016a)、統語的洗練性・複雑性分析のTAASSC (the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity) (Kyle, 2016; Lu, 2010) を使用した。また、ライティングスコアの代替と仮定して、Microsoft Word のファイルに対して計測できる文章の読みやすさを評価する指標、Flesch Reading Ease を使用した。

上記の方法で得られた指標について、経時的な比較等を行う際には、Social Science Statistics のオンラインツール、One-way ANOVA Calculator, Including Tukey HSD を使用した。

上記の方法で得られた指標とスコアの代替の指標との重回帰分析には、Statistics Kingdom のオンラインツール、Multiple Linear Regression Calculator を使用した。

### 3.3 分析手順

#### 3.3.1 一元分散分析

TAALES、TAACO、TAASSC のテキスト分析の後、過去の関連研究の結果を参照して、主要な指標を執筆時期やジャンル別に抽出した。

まず執筆時期について、41 本の作品を執筆順にほぼ同数 (11, 10, 10, 10) の 4 期間に分けて、一元分散分析 (one-way analysis of variance (ANOVA)) を行った。しかし、この区分は「学習者」の心理的な学習段階と必ずしも一致していなかった。また、最後の 3 作品の最初のもので、実際の試験時に書いたものを想起して書いたのに対して、残りの 2 本は (一次) 試験終了後に、合格をある程度、確信して、あまり明確な動機を持って執筆していなかった。このような理由から、最後の 2 本を除いて 39 本に絞った 3 期区分での分析も行うこととした。本研究では、二元分散分析は行っていないため、以下、ANOVA と呼称する。

ジャンル別分析については、「学習者」が分類したジャンル区分と各ジャンルの一般的な難易度が分析の基準となった。ジャンルは、①教育、②生活、③ビジネス、④政策、⑤環境、⑥科学に分かれ、「学習者」は、全体的な相対的難易度を①易しい、②&③やや易しい、④やや難しい、⑤&⑥難しい、と評価した。各ジャンルの本数は、① 3、② 16、③ 8、④ 6、⑤ 4、⑥ 4、と大きく異なったが、できるだけまとまった数の作品を比較するため、まず、①教育の 3 本は対象から外し、②、③からほぼ無作為 (1 本目の執筆作品をできるだけ優先し、連続的に執筆した同じテーマの 2 作品は避ける) に 6 本を抽出した。⑤、⑥は類似のトピックと考えて、3 本ずつ抽出して、共通ジャンル (環境・科学) としてまとめ、④の 6 本と、②、③、⑤+⑥の 6 本ずつの 4 ジャンルを比較することとした。その結果、②のように 6 本中、5 本が 1 本目の作品だったものもあれば、④のように、あるテーマの 1 本目から 4 本目をすべて含まなければならないものもあったが、あくまでも一つの参考比較として、ANOVA を行った。

なお、単一被験者の ANOVA 使用の賛否や方法については、様々な議論があるが、個人学習者の結果を英語学習者の特徴として一般化することが目的でないことと、エッセイのジャンルの比較に限らず、各期のエッセイの構成にも特段の対応関係がない ((例) 第 1 期 1 本目: 予想問題・生活・執筆 1 回目、第 2 期 1 本目: 過去問題・政策・執筆 3 回目、第 3 期 1 本目: 過去問題・生活・執筆 1 回目、第 4 期 1 本目: オリジナル問題・環境・執筆 1 回目) ことから、stats.stackexchange.com の議論等を参考にして、すべて、One-way ANOVA for independent measures (対応関係のない一元分散分析) を行い、F 値に有意差が確認された場合に、Tukey HSD (Tukey の HSD 法による多重比較) によって、すべての可能な平均の差を比較した。

#### 3.3.2 重回帰分析

3.3.1 の ANOVA で有意差が確認できた指標や過去の研究で議論されている指標とライティングスコアの代替として仮定した指標との関係を探るため、41 本のエッセイデータ

を対象にして、重回帰分析(ステップワイズ法)を行った。スコアの代替指標としては、リーダビリティの指標として広く使用される Flesch Reading Ease (FRE) について、まず検証を行った。

FRE を代替指標に選んだのは、FRE の数値が低くなるにつれて、文章の複雑さが増し、より高度な文章だと一般に見なされ、英語教育においても知名度が高いことが主な要因であるが、4 期区分や 3 期区分の ANOVA の結果も参考にした。

FRE 以外の指標についても目標変数に設定して、同様の分析を行った。TAALES、TAACO、TAASSC の指標の中から、4 期区分や 3 期区分の ANOVA や、FRE や他の多くの指標との単相関の数値を基準に、MLC (Mean Length of Clause) (節内の平均語数) を選定した。

各エッセイを英検の実際の基準で採点した場合に、FRE や MLC がどの程度その基準に関連するかは不明であるが、スコアの代替指標として仮定した場合に、スコアが付与されていないエッセイについても重回帰式から、スコアを予測することができるならば、実践的応用の指針となると考えた。しかしながら、複数の説明変数(独立変数)と単一の目標変数(従属変数)の関係は、あくまで予測であって、因果関係を説明するものではない。

### 3.3.3 質的分析

質的分析では、3 期区分の ANOVA の結果を、「学習者」が本実験に提供した実際のライティング、「学習者」がライティングに言及した日記とメモ用紙、「学習者」と指導教官との間で交わされたライティングに関する電子メール等と照らし合わせ、議論していく。

具体的には、3 期区分の ANOVA で有意差を示した指標を実際のライティングを例に出しながら、「学習者」のライティングがどのように変容していったかを分析する。

4 期区分と 3 期区分の ANOVA のうち、3 期区分のみを質的分析に用いるのは、3 期で分けることで、「学習者」の序盤・中盤・終盤のライティングを明確に区別できると考えたからである。

## 4. 数量的分析の結果

### 4.1 ANOVA

経時的な分析を行った先行研究としては、Crossley & Skalicky (2019) や Salsbury, et al. (2011) のスピーキングの分析で、Coh-Metrix や TAALES の指標である具体性、心像性、有意味性の数値が下がる傾向が報告されているが、本研究では、3 つのいずれの指標からも、同じ研究で調査された他の指標の親密性についても、有意差は確認されなかった。

次に、TAALES、TAACO、TAASSC の指標には含まれないが、本研究のデータ分析者が、過去の教育や研究の経験から、何らかの経時的な推移が見られるのではないかと考えて、FRE を分析対象に選んだ。表 1 は 4 期区分の結果 (41 本のエッセイ)、表 2 は 3 期区分の結果 (39 本のエッセイ) をまとめたものである。



表 1. FRE の平均値の推移 (4 期区分)

	1 期 (T1)	2 期 (T2)	3 期 (T3)	4 期 (T4)	F 値
エッセイ数	11	10	10	10	6.49 * $p < .05$
平均	63.8	55.8	61.7	50.0	Tukey HSD
SD	6.4	8.1	4.4	11.0	*T1>T4; T3>T4

表 2. FRE の平均値の推移 (3 期区分)

	1 期 (T1)	2 期 (T2)	3 期 (T3)	F 値
エッセイ数	13	13	13	4.44 * $p < .05$
平均	63.6	58.4	55.8	Tukey HSD
SD	6.5	8.1	5.6	*T1>T3

表 1、表 2 の結果から、4 期区分、3 期区分の分析ともに、FRE の平均値の推移は有意差を示し、Tukey HSD が 4 期区分では T1 と T4、T3 と T4、3 期区分では T1 と T3 の差が有意であることを伝えているが、4 期区分の 3 期でいったん数値が上がり（文章が易しくなり）、4 期で数値が下がっている（文章が難しくなっている）ことから一様な変化ではなかったことが推察される。FRE が実際の英検のライティング評価の基準に近似するものであるとすれば、2 期から 3 期にスコアがいったん下がったものの、3 期から 4 期にスコアが上がった可能性が示唆される。

FRE と同じような「一時停滞型」の変化を示した指標としては、TAALES の中に、Kuperman Age of Acquisition (AoA) (使用語単語の習得年齢平均値) と OLD (平均で何文字の字形変換で別の単語になるかの指標 <たとえば condition は 1, 2 文字の変化でたくさんの別の単語になる可能性があるが、pistachio は 4 文字変換してようやく別の単語になる特殊な綴りの単語 >) が確認された。FRE は 1 文の平均語数 (統語的洗練性に近い指標) と 1 語の平均音節数 (語彙洗練性に近い指標) を変数に含むが、他の 2 つの指標はいずれも単語の洗練性に関するものである。

一方で、4 期区分でも 3 期区分でも一様な変化を示した指標として、TAASSC の指標の中に、av\_nominal\_deps (dependents per nominal) (従属語の数/各名詞句)、MLC (節内の平均語数) が見られた。こちらは統語的な洗練性・複雑性の指標である。表 3 と表 4 は、MLC の 4 期区分と 3 期区分の平均値の変化を示したものである。

表 3. MLC の平均値の推移 (4 期区分)

	1 期 (T1)	2 期 (T2)	3 期 (T3)	4 期 (T4)	F 値
エッセイ数	11	10	10	10	7.88 * $p < .05$
平均	8.0	9.4	9.9	10.9	Tukey HSD
SD	0.9	1.4	1.0	1.9	*T4>T1; T3>T1

表 4. MLC の平均値の推移 (3 期区分)

	1 期 (T1)	2 期 (T2)	3 期 (T3)	F 値
エッセイ数	13	13	13	12.24 * $p < .05$
平均	8.1	9.8	10.1	Tukey HSD
SD	1.0	1.1	1.2	*T3>T1; T2>T1

実際の英検のライティングスコアの推移が、FREのような一時停滞型の変化だったのか、MLCのような様な進展型だったのかはわからないが、3期区分の平均値の推移で有意差を示したのものの中には他に、LD\_Mean\_RT (Lexical Decision Mean Reaction Time) (使用語彙を認識する平均反応時間)、simple subordinators (文章の語数全体に占める単純接続語の数 (1語で構成される接続表現))、節の数/文の数、次のパラグラフで使われる同じタイプの名詞 (type や types 等) や同義語 (type や kind 等) の比率、次の2つのパラグラフで使用される同じタイプの代名詞の比率が含まれ、語彙洗練性、文法洗練性・複雑性、テキスト結束性等、多様な側面において、ライティング機能に変化が生じていたことが示唆される結果だと言える。

一方、ジャンル別の比較においては、ANOVA で有意を示した指標のうち、「音声的に近い単語の数」を示す指標が②生活と③ビジネスの間で平均値に有意差を示したのに対して、「少なくとも次の2つの文で1回以上使用されるタイプの機能語の数の割合」が②と③と②と⑤+⑥ (環境・科学) の間 (表6参照)、「少なくとも次のパラグラフで1回以上使われるタイプの機能語の数の割合」が、②と③、②と④ (政策)、②と⑤+⑥の間、「少なくとも次の文で1回以上使用されるタイプの動詞の数の割合」が④と⑤+⑥の間、「少なくとも次の2つのパラグラフで1回以上使われるタイプの語の数の割合」が②と⑤+⑥の間、「3人代名詞の数/名詞の数」が、④と⑤+⑥の間において、Tukey HSD で平均値に有意差を示した。元々のジャンル別のエッセイの数が均等でないため、数量的なジャンル分析の結果を基に、過度の推量を行うことは避けるべきであるが、語彙レベルよりも文やパラグラフ間のつながりにおいて、ジャンル間の差異が顕著なようである。

表5. 「\*少なくとも次の2つの文で1回以上使用されるタイプの機能語の数の割合」の平均値の差異 (ジャンル別)

	②生活 (T1)	③ビジネス (T2)	④政策 (T3)	⑤ + ⑥環境・ 科学 (T4)	F 値
エッセイ数	6	6	6	6	5.57 * $p < .05$
平均	1.3	2.2	1.5	2.2	Tukey HSD
SD	0.5	0.5	0.5	0.2	* $T4 > T1$ ; $T2 > T1$

\*the number of function lemma types that occur at least once in the next two sentences ÷ number of sentences in text (except last two sentences) (文章中の文の数 (最後の2文を除く))

前置詞や代名詞等は目立たないが、文章の論理構造を支える重要な役割を担っている。語彙の洗練性や単文の構造の複雑さと同様に、機能語の使用状況がエッセイのジャンルによって差異が出ていることは、教授者や学習者にとっても注目に値する結果だと言える。

## 4.2 重回帰分析

仮に41本のエッセイの30本にスコアが付与されていても、残り11本のエッセイのスコアを文章の外観のみから客観的、かつ正確に予測することは容易にはできないだろう。しかしながら、オンライン分析ツールの様々な指標を併用することで、未採点のエッセイのスコアを予測することができる。本研究では、すべてのエッセイが未採点であるため、

スコアの代わりとして、文章の読みやすさや複雑さを示すとされ、Microsoft Word のファイルについて、簡単に計測できる Flesch Reading Ease (FRE) と呼ばれる指標を重回帰分析の目的変数（従属変数）に位置付け、TAALES, TAACO, TAASSC の様々な指標をステップワイズ法で検証し、最も効果的な予測が可能な説明変数（独立変数）の組み合わせを探った。

まず過去の研究で分析されている指標を中心に探したが、決定係数（寄与率）と呼ばれる相関値の 2 乗 ( $R^2$ ) が、過去の研究同様に、それほど高くならなかったため、FRE との単相関が相応の高さを示す指標についても確認した。他の説明変数と共分散性 (collinearity) の問題を生じない変数の選択を心掛けたが、問題が生じる場合は、分析ツールによって、自動的に変数から削除された。また、決定係数が高くなっても、残差の正規性 (residual normality) や等分散性 (homoscedasticity) について警告が出た場合は、候補から外した。

その結果、下記の重回帰式（数式 1）が得られた。

$$\text{FRE} = 206.698417 - 1.698447 \times \text{MLC} - 57.959427 \times \text{OLD} - 434.436773 \times 1/\text{Phono\_N} \quad (\text{数式 1})$$

\*MLC (Mean Length of Clause (節内の平均語数) )

\*OLD (綴り字の距離 (遠さ) を示す指標)

\*Phono\_N (音的に近い語の数の指標)

数式 1 の目的変数と説明変数の関係は、重相関値 (R) が .908、決定係数 ( $R^2$ ) が .824 (説明変数が目的変数の分散の 82.4% を説明することを意味する)、修正決定係数が、.810 と高い値を示した。ちなみに目標変数の FRE と各説明変数との単相関の値は、MLC が -.69、OLD が -.83、Phono\_N が .70 であったので、少なくとも、重回帰分析によって、より説明力の高い説明変数の組み合わせを見出すことができたと言える。

FRE は知名度も高く、本研究でも安定した指標として機能していたと言えるが、4.1 節の ANOVA の結果から、4 期区分で比較した時には、2 期から 3 期にかけて数値がいったん高くなっていた (文章の構造が容易になった) ことが、やや気付きである。実際に採点が行われた場合に、もし 4 期間に一樣な進展を示すスコアが得られることが予測されるならば、別の指標を目標変数に設定して、重回帰分析を行う必要がある。本研究では、実際のライティング能力がどのように推移していたかを断定できないため、4.1 節の ANOVA において、4 期間で安定的に一樣な変化を示した MLC を目標変数とする検証も行うこととした。

その結果、次のような重回帰式（数式 2）が得られた。

$$\text{MLC} = -14.688448 + 183.023511 \times 1/\text{OLDF\_CW} + 4.355482 \times \text{CP\_C} + 0.166495 \times \text{np\_elaboration} \quad (\text{数式 2})$$

\***OLDF\_CW** (綴り字が最も近い語 (内容語) の **HAL** (Hyperspace Analogue to Language) コーパスを基準にしたときの使用頻度の対数値の平均。語形の近さはその語を何文字入れ替え、挿入、削除すれば別の単語になるかで最も語形の近い 20 語を探し、その分布密度を測定する。たとえば、**condition** は 1 文字の挿入で **conditions**、2 文字の入れ替えで **coalition**、**cognition** 等があり、2 文字の挿入で **conditional** 等があり、3 文字までの変形で 20 語を見つけることができるが、**pistachio** は 4 文字の変換でようやく一番字形の近い **distraction** 等が得られる。)

\***CP\_C** (接続語句の数÷文章の節の数)

\***np\_elaboration** (名詞句の修飾関係の複雑さを示す複合的な指標。Kyle (2016) によると、19 の他の名詞句の指標と強い関連性があるとされる。)

数式 2 の目的変数と説明変数の関係は、重相関値 (**R**) が .911、決定係数 ( $R^2$ ) が .830 (説明変数が目的変数の分散の 83.0% を説明することを意味する)、修正決定係数が、.817 と高い値を示した。目標変数と説明変数の単相関の値は、**OLDF\_CW** が -.72、**CP\_C** が .50、**np\_elaboration** が .79 で、こちらも重回帰分析分析によって、より説明力の高い、効果的な説明変数の組み合わせを見出すことができたと言える。

数式 2 は決定係数の値も高く、ライティングスコア (代替値) を予測する効果においては、申し分ないが、数式の重要な構成要素である **np\_elaboration** の複合的な概念や数値が表す意味を理解することは、学習者や教師にとって容易ではないだろう。たとえば、**np\_elaboration** は ANOVA の平均値の分析においても、4 期区分、3 期区分ともに連続的な変化を示しているが、4 期区分の 1 期、2 期でマイナスを示した数値が、3 期、4 期でプラスを示していることが、何を意味しているのか、統計的にも、質的にも理解して、教育や学習の改善につなげていくことは難しい。そこで、**np\_elaboration** との相関が .92 を示した **av\_nominal\_deps** (名詞句における従属語の数の割合) を代替の説明変数としたところ、下記のような重回帰式 (数式 3) が得られた。

$$MLC = - 20.716332 + 195.828391 \times 1/OLDF\_CW + 4.775038 \times CP\_C + 4.340351 \times av\_nominal\_deps \quad (\text{数式 3})$$

数式 3 の目的変数と説明変数の関係は、重相関値 (**R**) が .890、決定係数 ( $R^2$ ) が .793 (説明変数が目的変数の分散の 79.3% を説明することを意味する)、修正決定係数が、.776 と、**np\_nominal** を説明変数に使用した時よりは若干数値が下がったが、かなり高い値を示した。目標変数と説明変数の単相関の値は、**OLDF\_CW** が -.72、**CP\_C** が .50、**av\_nominal\_deps** が .73 だった。重回帰分析から得られた数値は、いずれの単相関の 2 乗の値よりも高く、効果的な説明変数の組み合わせを見出すことができたと言える。

**av\_nominal\_deps** の計算方法が理解できれば、常にプラスの値で推移するこの指標のほうが学習者にとっても理解や点検がしやすいかもしれない。

## 5. 考察

### 5.1 質的分析

ここでは、3期区分のANOVAの結果をもとに「学習者」のライティングがどのように変わったかを分析していく。

まず、有意差を示した指標のうち、文章の読みやすさ（リーダビリティ）と語彙の洗練性に関する指標を4つ紹介する。1つ目は、Flesch Reading Ease Scoreで、代表的なリーダビリティの指標である。音節数が多い単語を使ったり、一つの文が長くなったりすると、その数値が低くなる。結果として、数値は下がる傾向にあり、1期と3期に有意差が見つかった。この結果から「学習者」は、1期から3期にかけ、使用する単語の音節数が増え、一つの文が長くなる傾向にあると推測できる。

2つ目に、Kuperman\_AoA\_AW (AoA) を挙げる。これは、ある単語における英語母語話者の習得にかかる年齢の指標の指標で、数値が上がる傾向にあった。FREと同様に、1期と3期に有意差が見つかった。この結果から、使用されている語における英語母語話者の習得年齢が上がっている傾向にあることがわかる。

3つ目の指標は、OLDFで、ある単語における使用頻度の指標である。数値が下がる傾向にあり、1期と3期に有意差が見つかった。これは、3期にかけて使用頻度の低い語を使った語が多くなることを示している。

最後にLD\_Mean\_RTを挙げる。これは、ある単語を識別するのにかかる時間を示し、数値が上がる傾向にあった。1期と3期に有意差があったことから、使用している単語を識別するのに時間がかかる文章が多い傾向にあることがわかる。

以下に挙げる例は、それぞれエッセイの本論から抜粋した文章である。斜体の下線部(例、*teenagers*)は、3音節以上の語を示している。例1と2からわかるように、3期の方が1期より、音節数の多い語を使って文章を書いている。例2の3期の例では、1, 2期の平均よりもFREとOLDFは低く、AoAとLD\_Mean\_RTが高くなっている。

例3は、FREの値が特に高くなった3期のエッセイの抜粋であるが、1, 2期の平均に比べて、OLDFの値が低く、AoAが高くなっている。curriculumやinteractiveのような音節数が多い教育関係のやや専門的な語を使用している点や、2つ目と3つ目の文が長いことが、文章全体のリーダビリティを下げて、関連する指標にも影響を与えていると思われる。

#### 例1

1期：学生のSNSの利用について書いたエッセイからの抜粋

First, I think *teenagers* can not use Social *Networking* Services safely because there are a lot of *dangerous* web sites such as web sites ...

3期：日常生活と音楽の関係性について（以下、略）

A *particular* sound or rhythm would make people *enthusiastically* or *negative*. Because of this, *companies* take *advantage* of that effect *creating commercials* with special sounds in order to ...

## 例 2

1 期：親元を離れて暮らす学生が増えていくかどうかについて

The another reason is that they may be leave their home because of desire of *personal* freedom. Since they have been watched by their parents for many years, they wants freedom lives, I think.

3 期：転職をする労働者が多くなるかどうかについて

The fact that *employees* quit jobs and seek others makes *companies* *difficult* to keep their *stability*. As a result, such *companies* are forced to employ more people *including* the *unemployed*.

## 例 3

3 期：最近の英語教育について

Another *important* issue is *curriculum*. Under the current teaching program, there are few *opportunities* for high school students to *communicate* with foreign students. Through a group-work or other *interactive* approaches, they need to *cooperate* with people who have *different* backgrounds.

次に、名詞句の修飾に関する指標を見ていく。av\_nominal\_deps は、名詞句の修飾表現の指標で、数値が上がる傾向が見つかり、1 期と 2 期及び 1 期と 3 期に有意差が確認できた。例 4 を見ると、1 期から 2 期、3 期と徐々に名詞を修飾する表現が増えている傾向にある（名詞とそれを修飾する表現は斜体下線部で示している）。初期の頃から長めの名詞句が使われているが、初期の段階とその後の時期の名詞句の修飾のし方を比較すると、学習者の成長を観察することができる。1 期では、such as A or B や in the number of ~ というような定型表現を用いて名詞を修飾している。それが、2 期になると定型表現にとられず、学習者自身の判断で、特に伝えたい名詞を関係代名詞を使って詳しく説明していることがわかる。3 期では、それがより顕著になる。名詞句を修飾する名詞句にも修飾表現が重なっており、全体で長い名詞句を形成している。また、名詞を修飾する関係節に関しても、節の中身が長くなる傾向になっていることが例からわかる。このことは、初期のライティングで限定的だった名詞句の修飾が、ライティングの練習を通して、様々な表現を学び、「学習者」が自由に名詞を詳しく説明することができるようになったことを示している。

## 例 4

1 期：環境保全について

First, *natural resources such as water, soil or something* are ... (中略) ... there is *the chance of increasing in the number of endangered species in Japan* ...

2 期：近い将来別の惑星に移住できるかどうかについて

There are too many things *that scientists have to study in the universe* ... (中略) ...

a lot of risks that they have never thought before.

3期：転職をする労働者が多くなるかどうかについて

The recent increase in the number of employees changing jobs has ...

同じトピックで書いた別のライティング

pressures on many companies that they need to improve job security ... (中略) ...

The fact that employees quit jobs and seek others makes ...

続いて、節に焦点を当てて検証をする。有意差が確認できた指標のうち、C\_SとMLCという指標について質的分析を行う。まず、C\_Sとは、「節の数÷文の数」で求められる指標で、数値は下がる傾向にあった。1期と3期に有意差が見られたことから、初期では頻繁に節を使っていたが、徐々にその頻度は減っていったことを示している。事実、1期でif節が2回以上使われているライティングがいくつか見つかった。その一方で、3期ではif節が使われているものの、その頻度は少ない。次に、MLCという指標について見る。これは、節内の平均語数を表し、数値は上がる傾向であった。また、1期と2期及び1期と3期に有意差が見つかった。

C\_SとMLCの結果からわかったことをまとめると、1期と3期にかけて節自体の頻度は少なくなっている一方で、節内の語数は増えている傾向がわかる。これには、120字から150字程度という制限語数がある英検ライティングの練習が関係していると考えられる。練習の初期段階では、制限語数に届くか届かない語数で文章を書いており、1期ではif節などの従属節を頻繁に用いることで文章を長く書こうと努力していた。しかし、徐々にその傾向が弱まっていき、3期になると無理やり節を増やすのではなく、その代わり具体的にメッセージを表現したいがために節内の語数が増え、自然と文章が長くなっていったことが推測される。具体的な例でその変化を見ていくことにする。

例5は、ともにif節を用いたライティングの抜粋だが、3期へ進むに従って節内の語数が増えている。また、2期まではif節の文が一つのライティングで頻繁に見られたが、3期になると一つのライティングにつき1回あるかないかの頻度になる。

例6は、that節に関する各期間の例である。節の部分は、斜体と下線で表してある。例5と同様、3期になるに従って節内の語数が増加しているのがわかる。3期に至っては、一つの文に同格のthat節と関係代名詞のwhichが存在し、長くて複雑な文になっている。このように、「学習者」のライティングの特徴を節に絞って見ると、1期から3期の期間に進むにつれ、節自体は少なくなる傾向が見られる一方で、一つ一つの節内の語数は多くなり、結果的に長くて複雑な文章が多くなることがわかる。

#### 例5

1期：歴史的建造物を保全すべきかについて

If the government did not do anything to protect historic sites, these sites would be destroyed.

If the places were not protected well, sightseers would feel disappointed.

2期：監視カメラを公共の場にもっと設置するべきかについて

... even if the government set more security cameras, our society could not be safer.

If the authority set more security cameras in public place, they cannot protect our privacy.

3期：環境保全について

For example, if the large area of our country were hit by severe drought, it would be hard to provide consumers with enough food supply.

例6

1期：環境保全について

For example, cutting a lot of tree cause a serious effect on creatures *that live in the forest*.

2期：監視カメラを公共の場にもっと設置するべきかについて

This contradicts with the original purpose *that the government provide comfortable facilities for inhabits*.

3期：転職をする労働者が多くなるかどうかについて

So, the recent trend *that people are switching jobs one by one* put a strain on a lot of companies *which are struggling with a shortage of workforce*.

最後に、名詞の繰り返しや言い換え表現（パラフレーズ）に関する指標について、`syn_overlap_para_noun` を説明する。この指標は、「次の段落で少なくとも1回以上使われる同じタイプの名詞や同義語の数÷最後のパラグラフを除いた段落の合計数」で求めることができる。数値は上がる傾向にあり、1期と3期に有意差が見つかった。この結果には繰り返しの語も多く含まれるが、初期には少なかった言い換え表現が3期に増加していることを示している。「学習者」にとっての言い換えは、よく使う語の繰り返しを避けるために用いる表現法で、同じ語を頻繁に繰り返すよりも、それを似た意味の内容でパラフレーズすることで、冗長さを無くし文章の読みやすさを追究する目的があった。これには、英検の二次面接（スピーキング）の練習時、瞬時のやり取りの中で、同じような内容を繰り返してしまいがちだが、それは無駄な発話と解釈され、良い評価に結びつかないことから、言い換えて話す訓練を意識したという背景がある。スピーキングで意識したことがライティングでも自然と意識され、言い換えの表現が増えたのではないと思われる。

具体的な例で言い換え表現について見ていく。例7の1期では、2つの段落に共通する名詞（斜体下線部）は2語で、そのうち *money* と *cash* は同義語のペアである。一方の3期は、共通の名詞が5語と数だけでも多く、そのうち *people* と *employees*、*company* と



firm、tendency と trend、security と stability が（文脈の中で）同義的に使われており、言い換えの表現が豊富である。

#### 例 7

1 期：日本がよりキャッシュレス化すべきかどうか（全体の 2 段落目と 3 段落目）

The one reason is that it would be very convenient for tourists to use money if there will be a completely cashless society. Thanks to a cashless system, they could exchange easily their own money into yen and they would be buying more products and using more services during a trip to Japan.

The other point is that it is easier for restaurants to do businesses. Because customer do not use cash when paying for something, cashier is not necessary. It could save much money and they could focus mainly on other sections such as service or food quality.

3 期：転職をする労働者が多くなるかどうかについて（全体の 2 段落目と 3 段落目）

Many people are eager to change jobs one to another because they have not been offered sufficient treatment by their companies. Although they often had been laid off before, the recent tendency puts so much pressures on many companies that they need to improve job security. Otherwise, such firms would be suffering from a shortage of workforce.

The recent trend also helps decrease unemployment rate in Japan. The fact that employees quit jobs and seek others makes companies difficult to keep their stability. As a result, such companies are forced to employ more people including the unemployed.

ここまで、「学習者」におけるライティングの変化を、数量的分析のデータから有意であった指標を基に、質的分析をしてきた。語彙的な面では、音節数が多い語が増え、一つの文が長くなり、リーダビリティの低い文章になった。また、名詞句を修飾する表現が増え、伝えたいことを詳しく説明できるようになった。節に関しては、節の数が減る一方で、節内の語数が増え、一文が長くて複雑なものになっていった。さらに、言い換えの表現が増えたこともわかった。「学習者」のライティングが、言いたいことを伝えるための語や表現が伴っていなかった初期に比べ、最終的には、伝えたいメッセージを自由に表現し、より複雑になったことがわかる。しかしながら、個人学習者のライティングの変化を多面的に考察するにあたり、未だ課題は残されている。例えば、説明のしやすさから、執筆の時系列ごとに 3 期に分けた ANOVA の結果を質的な分析の参考にしたため、「学習者」の微妙なライティングの変化を捉えることが難しかった点が挙げられる。また、数量的分析で得られた結果が、そのまま実際のライティングに反映されているとは限らず、その時の「学習者」の心理状態やトピックの書きやすさ等の諸要因が影響していることが考えられる。

## 5.2 今後の研究指針

本研究から得られた分析結果は、ある英語学習者個人の特徴を説明するものであり、日本の英語学習者や大学の英語学習者全般、あるいは異なる環境で学習する個人学習者に対

して、還元できる有用な情報は限られているかもしれない。しかしながら、ある英語学習者集団（サンプル）に対して行った分析結果が、その母集団に一般化できる情報を多く有していても、それが分析対象となったすべての学習者個人に適用できるわけではない。クラスを指導する教授者にとって、集団に準拠した研究から得られる知見は決して少なくないが、個別学習者の多様なニーズに具体的に対応していくことは、容易ではないだろう。個人学習者の技能習得の状況に関する詳細な量的及び質的な情報を提供することの意義は、英語教育・学習の研究において、より強く認識される必要がある。

本研究で分析の対象となった41本のライティング作品は、研究のために書かれたものではなく、学習者が英検準1級取得を目指して、純粋にライティング技能を磨くため、学習の目的で書いたものである。同じトピックに対して繰り返し書かれたものが大半を占め、そのうちの多くは、短い期間に連続的に作成されている。このような状況をふまえて行われた個人学習者の研究に、通常は集団の分析に使用されるANOVAを活用したことでなく、対応のない分析を行ったことの正当性には異論もあろうが、少なくとも多くの指標において、3期間、4期間区分の経時的比較の平均値に大きな差が見られ、ジャンル別分析においても、特徴的な差異が観測されたことは、事実である。また実際のライティング得点との比較ではないが、Flesch Reading Ease (FRE) のリーダビリティのスコアとMean Length of Clause (MLC) の統語的複雑性の指標を得点と見立てて、多くの指標との関係性を、重回帰分析から探求することもできた。今後は、他分野の単一被験者内の事象の変化における分析法等も参考にして、より明確な基準に基づいた数量的な検証を推し進めていくことが望まれるであろう。

今回の研究で活用したTAALES、TAACO、TAASSCのツールからは、800近い指標が得られたが、そのすべてを検証することはできなかった。過去の関連研究や本研究者の知見等を基に、有力な指標をできる限り多く抽出し、分析を行ったが、見落とされていた重要な指標があった可能性はある。また、計算過程や概念があいまいで、実践的なフィードバックに適さなかったり、質的な分析の結果に必ずしも結びつかない結果も見られた。

今後、このようなオンライン分析ツールはさらに開発が進み、上記の問題の多くが解消される可能性もあるが、分析者としても、より客観的で包括的、かつ実践的な分析方法を構築していく必要がある。

本研究では、質的研究は、主として数量的研究の結果を受けて行われた。今後は、ライティング執筆時に残されたメモ等を活用して、質的分析が数量的分析の基盤を築く可能性も視野に入れた研究を追究していくことが求められるだろう。

## 6. おわりに

本研究では、一人の学習者が英検準1級取得にむけて執筆したライティングをオンラインの分析ツールを用いて数量的に分析し、その結果から学習者のライティングがどのように変化したのかを考察した。数量的な分析で有意になったライティングの指標を、学習者のライティングや当時のことを言及した日記等と照らし合わせて質的に分析したところ、文章の読みやすさ（リーダビリティ）が低下していき、文章の複雑さが増したこと、名詞句の修飾表現や言い換え（パラフレーズ）表現の種類が豊富になったことが明らかになっ

た。ライティングをオンラインで分析することに関して、計算工程が不明な指標がある等、まだまだ議論の余地がある一方で、個人のライティングの変化を多角的に考察する手掛かりになったことは大きな成果だと考えられる。この分野の研究が進み、各指標の計算方法が明確になれば、学習者や指導者が個人でライティングを点検し、改善する指針となり得るだろう。

## 参考文献

- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016a). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, 48(4), 1227-1237.
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016b). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1-16.
- Crossley, S. A., & Skalicky, S. (2019). Examining lexical development in second language learners: An approximate replication of Salsbury, Crossley & McNamara (2011). *Language Teaching*, 52(3), 385-405.
- Hamada, A. (2015). Linguistic variables determining the difficulty of Eiken reading passages. *JLTA Journal*, 18, 57-77.
- Jung, Y., Crossley, S., & McNamara, D. (2019). Predicting second language writing proficiency in learner texts using computational tools. *Journal of Asia TEFL*, 16(1), 37-52.
- Kyle, K. (2016). Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication (Doctoral Dissertation).
- Kyle, K. & Crossley, S.A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly*, 49(4), 757-786.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4):474-496.
- Meara, P. (1995). Single-subject studies of lexical acquisition. *Second Language Research*, 11(2), i-iii.
- Salsbury, T., Crossley, S. A., & McNamara, D. S. (2011). Psycholinguistic word information in second language oral discourse. *Second Language Research*, 27(3), 343-360.
- 季駿. (2021, 3月). 「日本人英語学習者における英語エッセイのスペルエラー分析」(口頭発表)『第20回日本第二言語習得学会 国際年次大会 (J-SLA2020) 設立20周年記念大会』.(Web開催).
- 小室竜也. (2020). ライティングタスク(技能独立型 vs. 統合型)が発表語彙とその測定に与える影響: TAALESによる語彙の洗練性分析を基に(第32回研究助成)--(英語能力テストに関する研究).『「英検」研究助成報告』, 32, 13-30.

佐々木大和.(2020). 自動英文解析ツールを用いた英作文採点の妥当性検証：Coh-Metrix と Text Inspector の指標に基づいて (第 32 回研究助成)-(英語能力テストに関する研究).『「英検」研究助成報告』, 32, 31-43.