

学位論文 博士 (システムデザイン・マネジメント学)

重回帰分析により抽出した評価の観点
に基づく自由英作文のルーブリックのデザイン

指導教員 前野隆司 教授

2016 年度

慶應義塾大学大学院

システムデザイン・マネジメント研究科

システムデザイン・マネジメント専攻

平林 健治

目次

1. 序論	5
1.1 本研究の背景	5
1.1.1 自由英作文とは	5
1.1.2 本研究の対象者	8
1.1.3 自由英作文の重要性	15
1.1.4 自由英作文の普及に対する課題	18
1.2 本研究の目的	20
1.2.1 自由英作文のルーブリックの定義	20
1.2.2 研究目的	22
1.2.3 研究の意義	23
1.3 本研究の全体構成	25
1.4 まとめ	26
2. 評価手法の観点から分類した自由英作文のルーブリック	27
2.1 自由英作文の3つの評価手法（全体的評価・分析的評価・特定要因の評価）	27
2.2 全体的評価によるルーブリック	27
2.2.1 馬場（2008）のルーブリック	27
2.2.2 ‘Independent WRITING Rubrics’	29
2.3 分析的評価によるルーブリック	32
2.3.1 ESL Composition Profile	32
2.3.2 Nakanishi（2006）のルーブリック	35
2.3.3 Pilot Rubric for Paragraph / Essay Writing	36
2.4 分析的評価での評価の観点に特定要因の評価指標を用いた ルーブリックの考え方	39
2.5 評価の観点となる特定要因の評価指標の選択の考え方	40
2.5.1 自由英作文の言語的特徴に関する先行研究	40
2.5.2 CEFR-J A2 レベル自由英作文の評価の観点となる特定 要因の評価指標候補	42
2.5.3 CEFR-J B2 および C1 レベルの自由英作文の評価の観点 となる特定要因の評価指標候補	45
2.6 まとめ	47
3. CEFR-J A2 レベルの自由英作文の評価の観定の提案	50
3.1 CEFR-J A2 レベルの自由英作文のリーダビリティによる分析	50
3.1.1 はじめに	50

3.1.2	研究の背景	50
3.1.3	研究の目的	51
3.1.4	方法	52
3.1.5	結果と考察	54
3.1.6	結論	56
3.2	CEFR-J A2 レベルの自由英作文の語彙の豊かさの観点による分析	57
3.2.1	はじめに	57
3.2.2	研究の背景	57
3.2.3	研究の目的	59
3.2.4	研究方法	59
3.2.5	結果と考察	62
3.2.6	結論	65
3.3	重回帰分析により抽出した CEFR-J A2 レベルの自由英作文の評価の観点	67
3.3.1	はじめに	67
3.3.2	研究の目的	67
3.3.3	方法	67
3.3.4	結果	70
3.3.5	考察	74
3.3.6	結論	76
3.4	まとめ	77
4.	CEFR-J B2 および C1 レベルの自由英作文の評価の観点の提案	79
4.1	CEFR-J B2 および C1 レベルの自由英作文のリーダビリティによる分析	79
4.1.1	はじめに	79
4.1.2	研究の背景	79
4.1.3	研究の目的	80
4.1.4	方法	81
4.1.5	結果と考察	82
4.1.6	結論	84
4.2	CEFR-J B2 および C1 レベルの自由英作文の語彙の豊かさの観点による分析	86
4.2.1	はじめに	86
4.2.2	研究の背景	86
4.2.3	研究の目的	87
4.2.4	研究方法	88
4.2.5	結果	89
4.2.6	考察	92

4.2.7	結論	94
4.3	重回帰分析により抽出した CEFR-J B2 および C1 レベルの自由英作文 の評価の観点	96
4.3.1	はじめに	96
4.3.2	研究の目的	96
4.3.3	方法	96
4.3.4	結果	98
4.3.5	考察	104
4.3.6	結論	106
4.4	まとめ	107
5.	特定要因の評価指標を用いた評価の観点に基づく自由英作文の ルーブリックの提案	109
5.1	特定要因の評価指標を用いた評価の観点に基づくルーブリック のデザイン方法	109
5.1.1	ルーブリックのデザインに関する考え方	109
5.1.2	本研究のルーブリックの新規性	110
5.2	CEFR-J A2 レベルの自由英作文のルーブリックの提案	112
5.2.1	はじめに	112
5.2.2	研究の目的	112
5.2.3	研究方法	112
5.2.4	CEFR-J A2 レベルの自由英作文のルーブリックのデザイン	113
5.2.5	CEFR-J A2 レベルの自由英作文のルーブリック	116
5.2.6	結論	118
5.3	CEFR-J B2 および C1 レベルの自由英作文ルーブリックの提案	120
5.3.1	はじめに	120
5.3.2	研究の目的	120
5.3.3	研究方法	120
5.3.4	CEFR-J B2 および C1 レベルの自由英作文のルーブ リックのデザイン	121
5.3.5	CEFR-J B2 および C1 レベルの自由英作文のルーブリック	125
5.3.6	結論	128
5.4	まとめ	129
6.	提案した自由英作文のルーブリックの使い方と検証	131
6.1	CEFR-J A2 レベルの自由英作文ルーブリックの使い方	131

6.2	CEFR-J A2 レベルの自由英作文のルーブリックの検証	132
6.3	CEFR-J B2 および C1 レベルの自由英作文のルーブリックの使い方	135
6.3.1	Model 1 の CEFR-J B2 および C1 レベルの自由英 作文のルーブリックの使い方	135
6.3.2	Model 2 の CEFR-J B2 および C1 レベルの自由英 作文のルーブリックの使い方	143
6.4	CEFR-J B2 および C1 レベルの自由英作文のルーブリックの検証	145
6.4.1	Model 1 の CEFR-J B2 および C1 レベルの自由英 作文のルーブリックの検証	145
6.4.2	Model 2 の CEFR-J B2 および C1 レベルの自由英 作文のルーブリックの検証	147
6.5	まとめ	150
7.	結論	152
7.1	本研究の結論	152
7.2	今後の課題	159
	資料	160
	業績一覧	170
	謝辞	171
	参考文献	172

1. 序論

1.1 本研究の背景

1.1.1 自由英作文とは

1.1.1.1 自由英作文の定義

自由英作文ということばを文字通りに解釈すれば、書き手が束縛を受けずに、自らの書きたいことを英語で書き付けると言えるであろう。しかし、この解釈では、必ずしも本研究における自由英作文を明確に言い表しているとは言えない。そこで、以下に、本研究の自由英作文の定義を明らかにする。

本研究における自由英作文は、Richards et al. (1992: 147)による free writing の項目の “students write as much as possible about a topic within a given time period (for example, 3 minutes) without stopping. The goal is to produce as much writing as possible without worrying about grammar or accuracy, in order to develop fluency in writing and to produce ideas which might be used in a subsequent writing task. [(自由英作文とは) 学生が与えられた時間内(例えば、3分)で、あるトピック(テーマ)に対してできるだけ多くの英文を記述するものである。その目的はライティングにおける流暢さ(この場合、情報性の豊かさ、すなわち、書く量を表わす)の向上やライティングという課題に伴うアイデアの産出をはかるために、英文法や(語彙などの)正確さを気にすることなく、できるだけ多くの英文の記述をさせることである。]”を基本とする。すなわち、この定義では、英文の正確さよりも、英文の情報量を重視し、書き手のアイデアに重点を置くとしている。

本研究では、後述するように自由英作文の書き手である研究対象を日本人の初級英語学習者と上級英語学習者としている。前者の初級英語学習者については、自由英作文の学習始動期(1.1.2.3. CEFR-Jの「書くこと」のCAN-DOディスクリプタ参照)と考えられ、自由英作文の上達にはより多くの英文を記述することが望ましいため、Richards et al. (1992)の英文の正確さよりも英語の情報量を重視するという定義は意味のあることであると解釈する。一方、後者の上級英語学習者の自由英作文については、自由英作文の要件として、自由英作文の読み手に対して、書き手の意図を正しく伝達することを重要視し、その英文の正確さも含めるべきと考えた。Richards et al. (1992)では、自由英作文を記述するには連続した一定の時間内に英文を書くという条件が付加されており、本研究ではその時間を初級英語学習者では20分^{注1)}、上級英語学習者では30分^{注2)}とした。

注1) 本研究で使用したベネッセコーポレーション主催の英語コミュニケーション能力テストであるGTEC for STUDENTSのライティングセクションの解答時間である。

注2) Test of English as a Foreign Language (TOEFL)のライティングセクションの評価基 Independent WRITING(本研究の自由英作文に相当)の解答時間である。

従って、以下のように自由英作文を定義する。

(1) 日本人初級英語学習者の自由英作文

与えられたテーマに対して、学習者（書き手）の考えや意見を構築し、英語の言語形式の誤りよりも内容を重視し、20分という時間内に英文テキストを構成した説明文もしくは論述文と定義した。この初級英語学習者は一般的な日本人の高校生に相当する英語力である。

(2) 日本人上級英語学習者の自由英作文

与えられたテーマに対して、学習者（書き手）の考えや意見を構築し、英語の言語形式に正確であり、内容を重視し、30分という時間内に英文テキストを構成した説明文もしくは論述文と定義した。

1.1.1.2 自由英作文のメカニズム

自由英作文というのは日本人の書き手にとっては自らの意見や考えを構築し、それを外国語である英語で記述するため、様々な「ライティングの要素」が複雑にからむメカニズムを有する言語活動と言える。この要素について先行研究を基に以下に記述する。

「ライティングの要素」に関して、Raimes(1983)は、図 1-1 のようなモデルを提示している。Raimes(1983)によれば、自由英作文のようなライティングは「1. 統語法、2. 文法、3. 機械的技術、4. パラグラフの構成、5. 語の選択、6. 書く目的、7. 読み手、8. 書き手のプロセス、9. 内容」の 9 つの要素から構成されているとしている。緑川(1994)は 1 から 5 は形式に関わる領域、6 から 9 はコミュニケーションのための意味創造に関わる領域としている。図 1-1 は、自由英作文を通じて、書き手の「アイデアを明確に流暢に効果的に伝えるには、9 つのライティングの要素が満たされていなければならないことを示している。これらの要素は本研究におけるルーブリック^{注3)}の評価の観点とも密接に関わることになる。

注3) 自由英作文を評価するには一定の基準が必要となる。その評価基準や成績尺度を規定する指針をルーブリック (rubric) という (Nishijima, et al, 2007)。本研究では、評価の観点と評価基準 (到達の度合い) によるマトリックス型の表による形式を採用した (表 1-7 ライティング評価のためのルーブリックの例参照)。

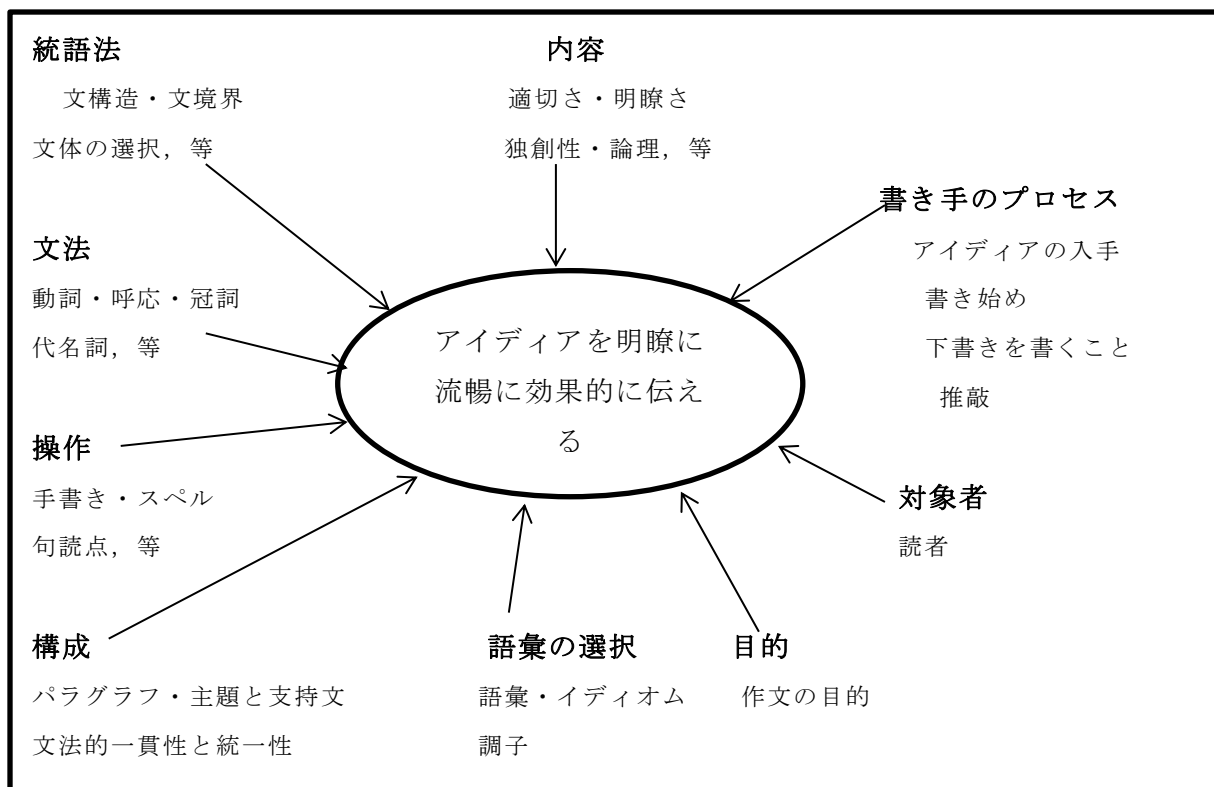


図 1-1 Raimes(1983)の 9 つのライティングの要素 (日本語訳は望月 (2010) から引用した)

次にこの要素の中の「書き手のプロセス」に関するメカニズムについて先行研究を基に以下に記述する。青谷 (2012 : 109-111) は、「書き手のプロセス」の視点から図 1-2 のようなモデルを提唱している。

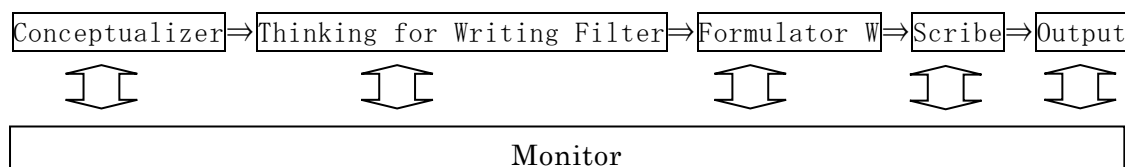


図 1-2 書き手のプロセスのモデル (青谷, 2012)

青谷 (2012 : 111) によると「Conceptualizer が作った意図を言語で表せる形にするため、単語を選択して spelling を含む諸情報を取得し、活用形を検討し、語順・文法・構文の処理を行ない、脳内に covert composition (内的作文) を形成する」とされる。このモデルによれば、書き手の意図を、「Thinking for Writing Filter」で言語的に思考し、それらを「Formulator W」で統合し、その結果を「Scribe」の段階で、文字化し英文が産出されると解釈される。そして、「Conceptualizer」から「Output」に至るすべてのプロセスにおいて、書き手は常に「Monitor」を通じて、自らのライティングプロセスを監視していることになる。

この「書き手のプロセス」を「プレライティング (pre-writing)」、「ホワイライティング (while-writing)」、「ポウストライティング (post-writing)」の3段階に分けた White and Arndt (1991) のモデルを以下に提示する。プレライティングでは、書く前の段階で、どのようなことを書くべきかを考え、「ホワイライティング」で英文を書き、「ポウストライティング」では推敲すると考えがちである。しかし、実際には、このような「プレライティング」⇒「ホワイライティング」⇒「ポウストライティング」という直線的な活動ではないことを White and Arndt (1991) は示している。図 1-3 が、ライティングのプロセスの各段階を進んだり、戻ったりしながら自由英作文が書き進められることを示している。

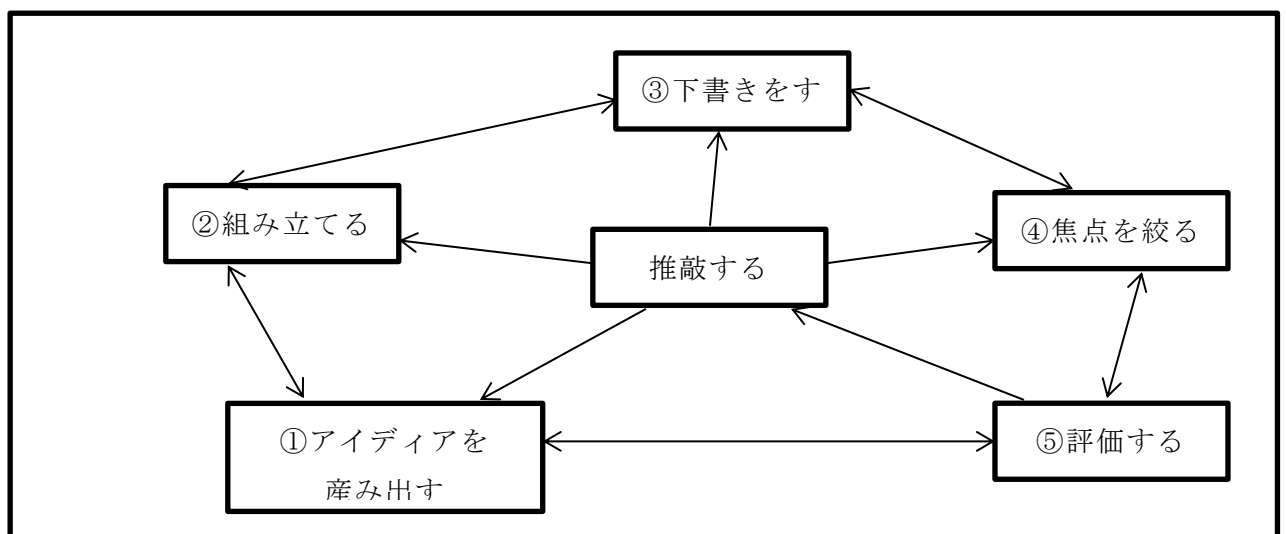


図 1-3 White and Arndt (1991) のライティングのプロセス (日本語訳は望月 (2010) から引用した)

このように自由英作文は、9 つのライティングの要素が統合され、ライティングのプロセスもアイディアの産出から推敲に至るまでの単純な直線的な進行ではなく、プロセスの各段階を前後しながら進行する非常に複雑なメカニズムによって成り立つ活動である。

1.1.2 本研究の対象者

1.1.2.1 本研究の対象者の抽出に対する考え方

本研究では、自由英作文の評価基準である既存のルーブリック (2 章参照) では、評価が困難となる日本人初級英語学習者と上級英語学習者を対象とする。初級英語学習者によって記述される自由英作文は、英語のライティング力の不十分さのために、その英文テキストの質は必ずしも高いものとは言えないため、既存のルーブリックでは適正な評価が困難になってしまうことになる。一方、外国人として英語を学ぶ完成段階にある上級英語学習者によって記述される自由英作文では、その英文の質の高さから既存のルーブリックで

は適正な評価が難しくなってしまうことになる。

このような初級英語学習者や上級英語学習者に関する定義付けとして、その英語力を一定の基準の中で捉えることとする。その基準として本研究では、欧州評議会によって作成された The Common European Framework of Reference for Language: Learning, Teaching, Assessment (CEFR) と呼ばれるヨーロッパにおける外国語学習、教育および評価のガイドラインに従うこととした。以下にこの CEFR という考え方を提示し、初級英語学習者と上級英語学習者の英語力を明確していくこととする。

1.1.2.2 CEFR-J という考え方

CEFR とは、欧州評議会が 2001 年に発行したヨーロッパにおける外国語学習、教育および評価のガイドラインである。この CEFR は言語共通参照枠として言語コミュニケーション能力を A (基礎段階の言語使用者)、B (自立した言語使用者)、C (熟達した言語使用者) の 3 つに分け、さらに 2 つの下位区分を設け、A1, A2, B1, B2, C1, C2 の 6 レベルに分類している。この 6 つのレベルに関して、「聞くこと」、「読むこと」、話すことの「やりとり」と「発表」、「書くこと」の 5 技能ごとに CAN-DO ディスクリプタと呼ばれる「～できる」という表現によってその能力を表示している (吉島他, 2004 ; 投野, 2013)。

CEFR-J は CEFR に準拠した日本版であり、2012 年 4 月に公開され、日本における新たな英語到達度指標として注目を集めている。Negishi, Takada, & Tono (2012) に示されているように日本人英語学習者の実態として 8 割が A レベルの初級に属すると言われる。そのため、CEFR-J では A レベルの下に Pre-A レベルを設定し、さらに A1, A2, B1, B2 レベルに関してもより細分化している。この CEFR と CEFR-J との関係を以下の表 1-1 CEFR と CEFR-J のコミュニケーションレベルに示した (投野, 2013)。

表 1-1. CEFR と CEFR-J のコミュニケーションレベル

CEFR		A1	A2	B1	B2	C1	C2
CEFR-J	Pre-A1	A1. 1	A2. 1	B1. 1	B2. 1	C1	C2
		A1. 2	A2. 2	B1. 2	B2. 2		
		A1. 3					

CEFR-J においても、「聞くこと」、「読むこと」、話すことの「やりとり」と「発表」、「書くこと」の 5 技能ごとに CAN-DO ディスクリプタと呼ばれる「～できる」という表現によってその能力が表示されている。この CAN-DO ディスクリプタは表 1-2 のようなマトリックス型の表で提示されている。マトリックスの各コラムに CAN-DO ディスクリプタが記載されている。例を示すと、表 1-2 の (a) の欄には「アルファベットの大文字・小文字、単語のつづりをブロック体で書くことができる」という Pre-A1 のライティングの CAN-DO ディスクリ

プタの1つが記載されている。

表 1-2. CEFR-J のコミュニケーションレベルの CAN-DO ディスクリプタのイメージ

	レベル	Pre-A1	A1	A2	B1	B2	C1	C2
理解	聞くこと							
	読むこと							
話すこと	やりとり							
	発表							
書くこと	書くこと	(a)						

次に、CEFR-J (CEFR) の A1～C2 で示される言語コミュニケーション能力と日本における英語の資格試験との関係を以下に示す。英語技能検定（以下、英検）の各級の英語力と CEFR-J (CEFR) の各レベルに関しては、日本英語検定協会によると以下の表 1-3 のようになる。

表 1-3. CEFR のレベルと英検のレベル

CEFR	対応する英検のレベル
C2	該当級なし
C1	英検 1 級（大学上級程度）
B2	英検準 1 級（大学中級程度）
B1	英検 2 級（高校卒業程度）
A2	英検準 2 級（高校中級程度）
A1	英検 3 級（中学卒業程度） 英検 4 級（中学中級程度） 英検 5 級（中学初級程度）

TOEIC と CEFR との関係については河原（2013：51-52）に「英検や NHK 英語番組のような細分化した形での試験形態ならば応用しやすいが、TOEIC といういわば一種類だけの試験に対しては応用は難しい。さらに TOEIC は読む、聞くという要素に特化した試験なので CEFR の 4 技能という分類に対して相応しづらい」と指摘しながらも、以下の表 1-4 のような対応表を提示している。

表 1-4. TOEIC と CEFR の対応表

Level	A1	A2	B1	B2	C1	Scale range
Listening	60	110	275	400	490	5-495
Reading	60	115	275	385	455	5-495
Total	120	225	550	785	945	

表 1-4 のように CEFR A1:TOEIC 120 点～225 点、CEFR A2:TOEIC 225 点～550 点、CEFR B1:TOEIC 550 点～785 点、CEFR B2:TOEIC 785 点～945 点、CEFR C1:TOEIC 945 点～990 点となる。このように TOEIC と CEFR の関係に着目すると、CEFR の各レベルを TOEIC の得点による基準で提示することができる。

本研究では英語力を一定の幅で把握することができるこの CEFR と TOEIC の関係に着目し、初級英語学習者および上級学習者の定義づけを行うこととする。

1.1.2.3 CEFR-J の「書くこと」の CAN-DO ディスクリプタ

CEFR-J の「書くこと」の CAN-DO ディスクリプタを表 1-5 に示した。この中で、本研究の自由英作文の定義に相当する CAN-DO ディスクリプタが現れ始めるのは A2 レベルである。A2.1-2(J) : 「文と文を and, but, because などの簡単な接続詞でつなげるような書き方であれば、基礎的・具体的な語彙、簡単な句や文を使った簡単な英語で、日記や写真、事物の説明文などのまとまりのある文章を書くことができる。」と A2.2-2(J) : 「聞いたり読んだりした内容（生活や文化の紹介などの説明や物語）であれば、基礎的な日常生活語彙や表現を用いて、感想や意見などを短く書くことができる。」の 2 つの CAN-DO ディスクリプタがそれに相当する。

A2 レベルよりも下位の英語力である Pre-A1 レベルや A1 レベルの CAN-DO ディスクリプタには本研究の自由英作文に相当すると考えられる記述はみられない。言い換えれば、CEFR-J では、「基礎段階の言語使用者」の上位者である A2 レベルの英語力が自由英作文の学習を始動する時期として、CAN-DO ディスクリプタを作成していると考えられる。表 1-4 によれば、この A2 レベルの英語力というのは、TOEIC 225 点～550 点であり、日本では、帰国子女を除く、大学進学を目指す一般的な高校生の英語力に相当することになる。

「自由英作文のメカニズム」で記述したように、自由英作文というのは、ライティングにおける 9 つの要素が絡み、記述の際にも複雑なプロセスを経ることが要求される。従って、自由英作文を記述するには一定の以上の英語力が必要となることになる。その英語力が CEFR-J A2 レベルであり TOEIC では 225 点～550 点ということになる。

CEFR-J B1 レベル (TOEIC 550 点～785 点) というのは、「自立した言語使用者」の下位レベルである。この B1 レベルの中での下位者に相当する英語力 (TOEIC 550 点～650 点程

度)は日本における一般的な大学生の英語力に相当する。また、B1 レベルの中での上位者 (TOEIC 730 点～785 点程度)は日本で TOEIC を主催している一般財団法人国際ビジネスコミュニケーション協会による「TOEIC のスコアとコミュニケーション能力レベルの相関表」の B レベル (TOEIC 730 点～860 点)に相当し、「どんな状況でも適切なコミュニケーションができる素地を備えている」という範疇に入る人たちの英語力となる。

(http://www.toEIC.or.jp/library/toEIC_data/toEIC/pdf/data/proficiency.pdf)

この B1 レベルにおける自由英作文の定義に相当する CAN-DO ディスクリプタは、B1.1-2(J) : 「身近な状況で使われる語彙・文法を用いれば、筋道を立てて、作業の手順などを示す説明文を書くことができる。」と B1.2-1(J) : 「新聞記事や映画などについて、専門的でない語彙や複雑でない文法構造を用いて、自分の意見を含めて、あらすじをまとめたり、基本的な内容を報告したりすることができる。」となっている。

B2 レベル (TOEIC 785 点～945 点) というのは、「自立した言語使用者」の上位レベルであり、日本では一般にかなり英語力が高いとみなされる人たちである。初等中等教育の英語教師や海外の大学や大学院への留学希望者、国際的なビジネスに従事する人たちの英語力に相当する。

この B2 レベルにおける自由英作文の定義に相当する CAN-DO ディスクリプタは、B2.1-2(J) : 「そのトピックについて何か自分が知っていれば、多くの情報源から統合して情報や議論を整理しながら、それに対する自分の考えの根拠を示しつつ、ある程度の結束性のあるエッセイやレポートなどを、幅広い語彙や複雑な文構造をある程度使って、書くことができる。」と B2.2-2(J) : 「感情や体験の微妙なニュアンスを表現するのであれば、重要点や補足事項の詳細を適切に強調しながら、筋道だった議論を展開しつつ、明瞭で結束性の高いエッセイやレポートなどを、幅広い語彙や複雑な文構造を用いて、書くことができる。」となっている。

C1 レベル (TOEIC®945 点～990 点) というのは、「熟達した言語使用者」の範疇に入り、日本では、高度な英語力を有すると見なされる人たちで、海外の大学院への留学希望者 (もしくは修了者)、国際的なビジネスに従事する人たち、プロの通訳者のレベルに英語力に相当する。

この C1 レベルにおける自由英作文の定義に相当する CAN-DO ディスクリプタは、C1(J) : 「いくつかの視点を示して、明瞭な構成で、かなり詳細に自己表現ができる。自分が重要だと思う点を強調しながら、手紙やエッセイ、レポートで複雑な主題について書くことができる。読者を念頭に置いて適切な文体を選択できる。」となっている。

最上位の C2 レベルというのは「熟達した言語使用者」の上位の英語力で、非母語話者として極めて英語が堪能な者となる。C2 レベルにおける自由英作文の定義に相当する CAN-DO ディスクリプタには「明瞭で流暢な文章を適切な文体で書くことができる。効果的な論理構造で事情を説明し、その重要点を読み手に気づかせ、記憶にとどめさせるよう、複雑な手紙、レポート、記事を書くことができる。仕事や文学作品の概要や評論を書くこ

とができる。」と記述されている。

表 1-5. CEFR-J の「書くこと」の CAN-DO ディスクリプタ

レベル	「書くこと」の CAN-DO ディスクリプタ
PreA1-1(J)	アルファベットの大文字・小文字、単語のつづりをブロック体で書くことができる。
PreA1-2(J)	単語のつづりを 1 文字ずつ発音されれば、聞いてそのとおりに書くことができる。また書いてあるものを写すことができる。
A1.1-1(J)	住所・氏名・職業などの項目がある表を埋めることができる。
A1.1-2(J)	自分について基本的な情報（名前、住所、家族など）を辞書を使えば短い句または文で書くことができる。
A1.2-1(J)	簡単な語や基礎的な表現を用いて、身近なこと（好き嫌い、家族、学校生活など）について短い文章を書くことができる。
A1.2-2(J)	簡単な語や基礎的な表現を用いて、身近な事柄についての簡単なハガキ、メッセージカード、短いメモなどを書ける。
A1.3-1(J)	自分の経験について、辞書を用いて、短い文章を書くことができる。
A1.3-2(J)	趣味や好き嫌いについて複数の文を用いて、簡単な語や基礎的な表現を使って書くことができる。
A2.1-1(J)	日常的・個人的な内容であれば、招待状、私的な手紙、メモ、メッセージなどを簡単な英語で書くことができる。
A2.1-2(J)	文と文を and, but, because などの簡単な接続詞でつなげるような書き方であれば、基礎的・具体的な語彙、簡単な句や文を使った簡単な英語で、日記や写真、事物の説明文などのまとまりのある文章を書くことができる。
A2.2-1(J)	身の回りの出来事や趣味、場所、仕事などについて、個人的経験や自分に直接必要のある領域での事柄であれば、簡単な描写ができる。
A2.2-2(J)	聞いたり読んだりした内容（生活や文化の紹介などの説明や物語）であれば、基礎的な日常生活語彙や表現を用いて、感想や意見などを短く書くことができる。
B1.1-1(J)	自分に直接関わりのある環境（学校、職場、地域など）での出来事を、身近な状況で使われる語彙・文法を用いて、ある程度まとまりのあるかたちで、描写することができる。
B1.1-2(J)	身近な状況で使われる語彙・文法を用いれば、筋道を立てて、作業の手順などを示す説明文を書くことができる。
B1.2-1(J)	新聞記事や映画などについて、専門的でない語彙や複雑でない文法構造を

	用いて、自分の意見を含めて、あらすじをまとめたり、基本的な内容を報告したりすることができる。
B1. 2-2 (J)	物事の順序に従って、旅行記や自分史、身近なエピソードなどの物語文を、いくつかのパラグラフで書くことができる。また、近況を詳しく伝える個人的な手紙を書くことができる。
B2. 1-1 (J)	自分の専門分野であれば、メールやファックス、ビジネス・レターなどのビジネス文書を、感情の度合いをある程度含め、かつ用途に合った適切な文体で、書くことができる。
B2. 1-2 (J)	そのトピックについて何か自分が知っていれば、多くの情報源から統合して情報や議論を整理しながら、それに対する自分の考えの根拠を示しつつ、ある程度の結束性のあるエッセイやレポートなどを、幅広い語彙や複雑な文構造をある程度使って、書くことができる。
B2. 2-1 (J)	自分の専門分野や関心のある事柄であれば、複雑な内容を含む報告書や論文などを、原因や結果、仮定的な状況も考慮しつつ、明瞭かつ詳細な文章で書くことができる。
B2. 2-2 (J)	感情や体験の微妙なニュアンスを表現するのでなければ、重要点や補足事項の詳細を適切に強調しながら、筋道だった議論を展開しつつ、明瞭で結束性の高いエッセイやレポートなどを、幅広い語彙や複雑な文構造を用いて、書くことができる。
C1 (J)	いくつかの視点を示して、明瞭な構成で、かなり詳細に自己表現ができる。自分が重要だと思う点を強調しながら、手紙やエッセイ、レポートで複雑な主題について書くことができる。読者を念頭に置いて適切な文体を選択できる。
C2 (J)	明瞭で流暢な文章を適切な文体で書くことができる。効果的な論理構造で事情を説明し、その重要点を読み手に気づかせ、記憶にとどめさせるよう、複雑な手紙、レポート、記事を書くことができる。仕事や文学作品の概要や評論を書くことができる。

1.1.2.3 日本人初級英語学習者：CEFR-J A2 レベル

本研究の対象とする初級英語学習者は自由英作文の学習を始動する時期である英語力の CEFR-J A2 レベル (TOEIC 225 点～550 点) の「基礎段階の言語使用者」の上位者とする。この英語力は上述したように、帰国子女を除く、大学進学を目指す一般的な高校生の英語力に相当すると考えられる。

表 1-3 の日本英語検定協会によるレベル提示によれば、英検準 2 級 (高校中級程度) となっている。この基準を TOEIC の得点が 225 点から 550 点までの 325 点もの差のある英語

力に一律に当てはめるのはやや問題があると考えた。そこで、筆者は TOEIC を国際ビジネスコミュニケーション協会のホームページに掲載されている TOEIC の得点(第 86 回公開テスト) と英検取得者の関係を確認してみた。結果として、英検 2 級・準 2 級・3 級の各級の取得者の最も多いのが、それぞれ 445 点～495 点、345 点～395 点、245 点～295 点ということが判明した。この結果を参考すると、CEFR-J A2 レベルを 3 段階に分け、上位者は英検 2 級程度、中位者は英検準 2 級程度、下位者は英検 3 級程度であると判断される。

CEFR-J A1 レベルの上位者が英検 3 級の範疇に入ることから、CEFR-J A2 レベルの下位者は英検 3 級の合格者の中でも上位に位置すると考えられる。そして、CEFR-J A2 レベルの CAN-DO ディスクリプタで示されている自由英作文相当の英文を記述するには、最低限、英検 3 級の上位者レベルの英語力を有している必要があると判断される。

資料 1-1 に CEFR-J A2 レベルの学習者によって記述された自由英作文の事例を提示した。

1.1.2.4. 日本人上級英語学習者：CEFR-J B2 および C1 レベル

本研究のもう一方の対象となる日本人上級英語学習者は、B2 レベルおよび C1 レベルとする。TOEIC では、785 点から満点の 990 点までの英語力を有する学習者となる。CEFR における B2 レベルは「自立した言語使用者」の上位レベルであり、C1 レベルは「熟達した言語使用者」となる。B2 レベルおよび C1 レベルというのは日本人として英語学習に成功した上級学習者と一般的に考えられている人達である。この英語力は上述したように、英語教師や海外の大学院への留学希望者もしくは修了者、国際的なビジネスに従事する人たちの英語力に相当すると考えられる。小池他(2010)の調査で、国際交渉の第一線で活躍するビジネスパーソンの 70%が TOEIC 800 点以上が必要であることをアンケート調査で示しているが、CEFR-J B2 および C1 レベルは、ほぼその範疇に入る英語力を有していることになる。

資料 1-2 に CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文の事例を提示した。

1.1.3. 自由英作文の重要性

1.1.3.1 CEFR-J A2 レベルの学習者にとっての自由英作文の重要性

国際化著しい社会の中で、自由英作文という発信型の英語教育活動の重要性は今後ますます高まると考えられる。CEFR-JA2 レベルの初級英語学習者が対象となる中央教育審議会の答申「新しい時代にふさわしい高大接続の実現に向けた高等学校教育、大学教育、大学入学者選抜の一体化改革」(平成 26 年 12 月公開)の中に「グローバル化の進展の中で、言語や文化が異なる人々と主体的に協働していくためには、国際共通語である英語の能力を、真に使える形で身に付けることが必要であり、単に受け身で「読むこと」「聞くこと」ができるというだけではなく、積極的に英語の技能を活用し、主体的に考えを表現することができるよう、「書くこと」「話すこと」も含めた四技能を総合的に育成・評価することが重

要である。」とあり、主体的に考えを表現するライティング活動の重要性が強調されている。ライティング活動の一翼を担うのが本研究の自由英作文ということになる。

どのような「学び」においても、当初の「学び」が、それを成功に導くために大切となる。自由英作文の場合、CEFR-JA2 レベルがその始動期であり、この時期に順調に自由英作文を学んだ学習者は、その後の学びが順調に推移する可能性が高まると言えよう。ここでは、この始動期における自由英作文の学びの重要性を (1) 高等学校学習指導要領、(2) 大学入学試験、(3) 英語の資格試験の視点から以下に記述する。

現行の高等学校学習指導要領は平成 25 年 4 月 1 日の入学生から年次進行により段階的に適用されているものである。この学習指導要領では「コミュニケーション英語 I, II, III」により 4 技能を統合した教授法が推奨され、ライティング活動もその中で、指導されることになる。本研究の自由英作文のような活動は、「英語を通じて、積極的にコミュニケーションを図ろうとする態度を育成するとともに、事実や意見などを多様な観点から考察し、論理の展開や表現の方法を工夫しながら伝える能力を養う。」を目標として新設された「英語表現 I, II」の中で培われることになる。このような科目が新設されたことは、国際化の著しい社会の中で、国際共通語になりつつある英語という言葉で、自由英作文の始動期である高校生の段階から、自己の考えを英語にして発信していくことの重要性を示している。

学習者指導要領の改訂に伴い、大学入学試験の英語の学力試験において自由英作文に相当する問題がどの程度出題されているのか、筆者は旺文社の『全国大学入試問題正解』を基に平成 26 年度の国公立大学 2 次試験を対象として調査してみた (資料 1-3 参照)。結果として、全国の国公立大学の 2 次試験の 81 の英語問題の中で、35 の問題 (43%) で、自由英作文に相当する問題が出題されていることが明らかとなった。高校生や受験生にとっては、かなり高度な自由英作文の記述を求めている大学 (例えば、一橋大学、浜松医科大学など) もあることもあわせて明らかとなった。但し、国公立大学で自由英作文を出題する大学の受験者については CEFR-J B1 レベルの英語力 (英検 2 級以上、TOEIC 550 点~785 点) の可能性があるので、CEFR-J A2 レベルの英語力を対象とする本研究の対象者とは必ずしも一致しないことも考慮しておく必要がある。

資格試験においては、CEFR-J A2 レベルに相当する高校生は TOEIC や TOEFL よりは英検の受験が一般的である。英検を主催する日本英語協会は、英語学習者の 4 技能のバランス良い習得を目的として、2016 年度 6 月実施の試験から、英検 2 級にライティング試験を導入することを発表している。

(http://www.eiken.or.jp/eiken/info/2015/pdf/20150715_pressrelease_writing2.pdf)

出題形式は以下のような問題例を示している。

- 以下の TOPIC について、あなたの意見と理由を 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 語～100 語です。

TOPIC

These days, some people buy things on the Internet. Do you think more people do so in the future?

POINTS

- Cost
- Safety
- Technology

この出題例は本研究で扱う自由英作文に相当する。英検 2 級は CEFR (CEFR-J) の B1 レベルを意識して作成されているが、「1.1.2.3 日本人初級英語学習者：CEFR-J A2 レベル」に記述したように、A2 レベルの上位者は英検 2 級に合格する英語力を有していると考えられるため、資格試験として英検 2 級に自由英作文が導入されたことによって、A2 レベルの学習者にとっても自由英作文の重要性が示されたことにもなる。

また、同協会によると、英検準 2 級および 3 級においてもライティングの問題の出題をできるだけ早急に実現することを検討しているとのことである。出題形式は英検 2 級に類似したものになると予想されるので、自由英作文が CEFR-J A2 レベルに相当する英検準 2 級および CEFR-J A1 レベルに相当する英検 3 級（但し、「日本人初級英語学習者：CEFR-J A2 レベル」に記述したように、CEFR-J A2 レベルの下位者は英検 3 級の英語力と重なる可能性がある）近い将来、導入されることになる。

このような高等学校学習指導要領や大学入試および資格試験の動向は、CEFR-J A2 レベルの学習者にとって、中央教育審議会の答申にある「主体的に考えを表現する」ことの重要性を如実に反映したものとなっていると言えよう。

1.1.3.2 CEFR-J B2 および C1 レベルの学習者にとっての自由英作文の重要性

CEFR-J B2 および C1 レベルの学習者は、その多くが国際的なビジネスや学術の分野で活躍する人達である。国際的なビジネスにおいては、日々の業務において、電子メールなどで、英語で業務に関する自身の考えを記述することは日常茶飯である。また、学術の分野で活躍する人達は、どの専門分野であろうと英語で論文やエッセイを記述することが求められる。このレベルの英語学習者にとって、自由英作文というのは必要不可欠なライティ

ング活動とすることができる。

このレベルの日本人英語学習者の資格試験として最もよく知られている英検 1 級 (CEFR-J C1 相当) では、「社会生活一般、芸術、文化、歴史、教育、科学、自然・環境、医療、テクノロジー、ビジネス、政治など」に関するテーマで自由英作文を記述する問題が出題されている。

(日本英語検定協会：http://www.eiken.or.jp/eiken/exam/grade_1/detail.html) この出題は、上級の英語学習者にとって、与えられたテーマに対して、自身の考えを英語で記述する自由英作文の必要性を求めていることになる。

また、欧米の大学や大学院への留学希望者は Test of English as a Foreign Language (以下、TOEFL) においても、そのライティングセクションにおいて、Integrated 型と Independent 型の 2 つタイプの課題が課される。前者は学術的トピックについての英文を読み、それに関連する 2 分の講義を聞き取り、それについて 20 分で英文を書き上げるものであり、後者は課題が与えられ、意見を 30 分で書き上げるものである。この後者の Independent 型の課題は本研究の自由英作文の定義とほぼ一致する。このような自由英作文能力は留学先である大学や大学院で求められるエッセイ作成能力を検査するには不可欠と判断されて出題されていると考えられる。

このように CEFR-J B2 および C1 レベルの学習者のような上級英語学習者にとっても自由英作文というライティング活動はビジネス、学術のいずれの分野においても重要な役割を果たしており、その能力を測定するために資格試験においても課題として出題されているのである。

1.1.4 自由英作文の普及に対する課題

自由英作文は「ライティングの要素」や「ライティングプロセス」において非常に複雑なメカニズムによって成し遂げられるライティング活動である。そのため、日本人の書き手が外国語である英語で与えられたテーマに対して自身の考えや意見を構築し自由英作文を書くということは困難な活動となりがちである。また、その書き手の指導者にとっても、自由英作文の指導は難しいと言われることが多い。本研究の初級英語学習者の指導者である高等学校の先生方から自由英作文について「指導方法がわからない。添削するのが大変であり、時間もない」という声があがり、パラグラフ単位で英文を書く活動が敬遠されることになる (馬場, 2010: 119)。日本、韓国、台湾の 3 か国の高校の先生方を対象として、ライティング指導に関する調査を行った大井 (2015) によれば、日本人の先生方の指摘で最も多かったのが、「教員のライティングに関する知識不足および指導力不足」という項であったとも報告されている。このことは、上級の学習者にとっても当てはまり、例えば、TOEFL における Independent Writing 型 (本研究の自由英作文に相当) において自身の書いた英文の自己採点が困難であり、その指導者もどの点を改善すべきかが明確にわかっていないという声もあがることになる (植田他, 2015)。結果的に、日本においては自由英作

文に対する重要性は高まっている一方、自由英作文というライティング活動の教育実践が困難に直面していることになる。

このような状況の大きな原因の1つは自由英作文に対する評価基準の指針が日本の英語教育において十分に整備されていないことにある。その評価基準の指針が本研究で扱うルーブリックと言われるものである。ルーブリックとして、最も汎用的なものとして、TOEFLのライティングセクションの Independent Writing 型の課題に対する ‘Independent WRITING Rubrics’ や Jacob et al(1981)の ESL Composition Profile が有名である。これらはアメリカで開発されたものである。また、日本の研究者によって作成されたルーブリックに馬場 (2008)、Nakanishi(2006)、Nishijima, et al(2007)がある。これらの既存の5つのルーブリックと本研究でデザインするルーブリックの特徴を表 1-6 にまとめた(既存のルーブリックの詳細は2章参照)。

表 1-6 5つの主なルーブリックと本研究でデザインするルーブリックの特徴

	馬場 (2008)	Nakanishi (2006)	Nishijima et al (2007)	Independent WRITING Rubrics	ESL Composition Profile	本研究	本研究
対象者 (ルーブリック作成の際の英語力の測定)	大学生 (英語力の測定なし)	大学生 (英語力の測定なし)	大学生 (英語力の測定なし)	英語学習者全般	英語学習者全般	CEFR-J A2 レベル (英語力測定あり)	CEFR-J B2 および C1 レベル (英語力測定あり)
評価手法	全体的評価	分析的評価	分析的評価	全体的評価	分析的評価	特定要因の評価の指標に基づく分析的評価	特定要因の評価の指標に基づく分析的評価
言語的特徴の実証研究の有無	無	無	無	無	無	有	有

表 1-6 のように、先行研究で扱われている日本人学習者向けの自由英作文としての作成されたルーブリックも 3 つしかなく、日本での自由英作文のルーブリックに関する研究はあまり進んでいるとは言えない状況である。また、これらのルーブリックは日本の EFL (English as a Foreign Language の略で、外国語としての英語という意味) 環境に見合った評価基準を提案しようとしているが、今後、妥当性を調査し、さらなる改良が望まれる (馬場, 2010) と指摘されているように、その質的な改善も望まれている。表 1-6 に提示した馬場 (2008)、Nakanishi (2006)、Nishijima et al (2007) のいずれのルーブリックでも、対象者の英語力は測定されておらず、その英文の言語的特徴を実証することなく、デザインされていることから、質的な改善の余地は大いにあると言えよう。また、大学生を対象としてのみルーブリックが作成されており、それ以外の日本人学習者の対象としたルーブリックの研究はほとんどなされていないのが現状である。日本における自由英作文の重要性はその学習の始動期に当たる CEFR-J A2 レベルの初級学習者や CEFR-J B2 および C1 レベルの上級者学習者にとっては高まっているにもかかわらず、それらの学習者を対象としたルーブリックが存在しないことになっているのである。

このような状況の中で、その汎用性の高さから ‘Independent WRITING Rubrics’ が日本で最も利用されるルーブリックの 1 つとなっているが、アメリカで開発されたため、EFL の環境の日本人学習者には必ずしも適していないという欠点がある。このルーブリックでは 0 点～5 点までの基準が示されているが、CEFR-J A2 レベルの自由英作文の評価は大半が 1 点の範疇にはいり、英文の質に差のある場合でも、評価の峻別ができなくなってしまう。CEFR-J A2 レベルは英語力において英検 3 級から 2 級の範疇に入ると考えられるので、その自由英作文能力にも相当の差があると考えられるが、それを峻別できず、‘Independent WRITING Rubrics’ は評価基準として機能しないことになる。また、CEFR-J B2 および C1 レベルの上級者によって記述された自由英作文に対する評価は通常、4 点もしくは 5 点となると想定されるが、評価を示す記述があまりに一般的かつ抽象的な表現が使われており (Nishijima, et al, 2007)、上級英語学習者の自由英作文の評価にも必ずしも有用とは言えないという欠点がある。

このように CEFR-J A2 レベルの初級者や CEFR-J B2 および C1 レベルの上級者によって記述される自由英作文を評価するためのルーブリックがデザインされれば、自由英作文の評価基準の指針が明確になることにより、自由英作文の指導者には評価が容易になるだけでなく、指導方法の確立にも有用となり、学習者にとっても学習の目安を提示されることになる。その結果として、自由英作文というライティング活動の教育実践がより普及することになると考える。

1.2 本研究の目的

1.2.1 自由英作文のルーブリックの定義

自由英作文を評価するには一定の基準が必要となる。その評価基準や成績尺度を規定す

る指針をルーブリック (rubric) という (Nishijima, et al, 2007)。研究社の『新英和大辞典 第六版』によれば、英語の ‘ rubric ’ には形容詞で「朱文字で書かれた」、また、動詞で「朱書きする」、名詞に「(法令・規定などの) 題目、項目」という意味があり、法令や規定の「題目」が昔は朱書きされたと記述されている。さらに、教会の「典礼法規」とも記載されている。このような原義を持つため、「評価基準や成績尺度を規定する指針」に対して、このルーブリックという単語が使われるようになったと考えられる。

ルーブリックは自由英作文の評価に使われるだけでなく、教育機関の評価指標として近年、日本においても注目されているものである。そのようなルーブリックについて、関西国際大学学長で中教審大学教育部会の委員を務める濱名篤先生の中教審大学教育部会 (2011 年 12 月 9 日) 説明資料を以下に提示す。

(http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/015/attach/1314260.htm)

- (1) 「目標に準拠した評価」のための「基準」つくりの方法論であり、学生が何を学習するのかを示す評価規準と学生が学習到達しているレベルを示す具体的な評価基準をマトリクス形式で示す評価指標である。(俣野秀典の作成した定義をもとに濱名が修正したもの)
- (2) 学習者の「パフォーマンスの成功の度合いを示す尺度と、それぞれの尺度に見られるパフォーマンスの特徴を説明する記述語で構成される、評価基準の記述形式」として定義される評価ツールのことであり、アメリカにおいて先進的に開発され、数多くの高等教育機関が導入・活用している。(吉田武大「アメリカにおけるバリュールーブリックの動向」関西国際大学教育研究所研究叢書 4 号、2011 年 3 月)
- (3) レベルの目安を数段階に分けて記述して、達成度を判断する基準を示すものである。学習結果のパフォーマンスレベルの目安を数段階に分けて記述して、学習の達成度を判断する基準を示す教育評価法として盛んに用いられるようになった。これまでの評価法は客観テストによるものが主流を占めていたが、知識・理解はそれで判断できたとしても、いわゆるパフォーマンス系 (思考・判断、スキルなど) の評価は難しい。ポートフォリオ評価などでルーブリックを用いて予め「評価軸」を示しておき、「何が評価されることがらなのか」についての情報を共有するねらいもある。(熊本大学 WEB 「学習指導・評価論」より)
- (4) 「ルーブリック評価指標」は「それぞれの尺度に見られるパフォーマンスの特徴を示した記述語 (descriptor)」と「達成の度合いを示す数値的な尺度 (scale)」で、評価指標を設定しようという考え。「記述語」というのは、評価の視点、或いは観点にあたるもので、「尺度」というのはレベル、段階的水準ともいえるもの。学校現場では、「評価規準」と「評価基準」という用語がよく使われているが、その用語の使用法と似ているところがある。(高浦勝義『絶対評価とルーブリックの理論と実践』黎明書房、2004 年より)

このようなルーブリックは日本における英語教育においても、活用され始めている（河内山他，2013；Brendan，2014）。しかしながら、英語教育のライティング、それも自由英作文に限定すると、その研究は少なく（馬場，2008；Nakanishi，2006；Nishijima et al, 2007 など）、活用も十分になされているとは言い難い状況である。

本研究で扱うルーブリックは上述した濱名先生の定義に準拠することにする。以下に、杉田（2015）^{注4)}が、ライティングの評価のためのルーブリックを解説するために使用した例をもとに本研究における自由英作文のルーブリックの用語等を定義する。杉田（2015）ではルーブリックを解説するために表 1-7 のようなルーブリックを示している。

表 1-7 ライティング評価のためのルーブリックの例

評価規準 (観点)	評価基準 (到達の度合い)		
	A	B	C
文の正確さ	現在完了の用法、単語のスペリングや句読法が正しい。 (6点)	現在完了の用法は正しいが単語のスペリングや句読法に誤りがある。 (4点)	現在完了の用法、単語のスペリングや句読法に誤りがある。 (2点)
内容の適切さ	150～200 語 (4点)	100～149 語 (3点)	100 語未満 (2点)

表 1-7 は濱田の説明の(1)～(4)の説明に準拠するルーブリックとなっている。本研究でデザインするルーブリックもこのようなマトリクス形式とする。学生が何を学習するのかを示す「評価規準」と学生が学習到達しているレベルを示す具体的な「評価基準」という用語に関して、「規準」と「基準」が日本語では同じ音になり、両者を鮮明に区別するために、前者の「評価規準」を本研究では「評価の観点」という用語で統一することにする。

1.2.2 研究目的

日本社会が直面する国際化は、英語教育に対して発信型の英語教育活動の実践をますます求めることになっている。発信型の英語教育活動の中でも近年、注目されているが4技能の内のライティング活動における、自身の考えや意見を英語にして記述する自由英作文である（平成26年12月公開の中央教育審議会の答申「新しい時代にふさわしい高大接続の実現に向けた高等学校教育、大学教育、大学入学者選抜の一体化改革」参照）。それは高

^{注4)} 杉田（2015）では、中学校や高等学校の英語教師を対象としてルーブリックを説明するために、表 1-7 を提示しているが、これは、は現在完了という英文法事項を生徒に学ばせるために、自由英作文を書かせた場合のルーブリックの例として作成されたものである。

等教育の学習指導要領で示される高等学校の英語の科目に反映し、大学入試における英語の問題にも自由英作文の出題される傾向が高まることにもなる。また、海外の大学や大学院に入学を希望したり、国際的なビジネスや学術の分野で活躍する上級の英語学習者にとっても、自身の考えを英語で記述することは非常に重要な能力となっている。

このようなニーズがあるにも関わらず、自由英作文の学習の普及が必ずしもすすんでいない傾向がある。その原因の1つに自由英作文の評価基準の指針であるルーブリックが十分に整備されていないことにあると考えられる。評価基準の指針が定まっていないため、学習者にとっては自由英作文に対する学びに支障をきたしがちになり、指導者も指導方法に戸惑うことになる。

そのような状況を是正するために、近年、日本でもルーブリックに関する研究が行われて（馬場，2008； Nakanishi, 2006； Nishijima et al, 2007）いるものの、対象が日本人の大学生のみとなっており、自由英作文の言語的特徴を明らかにしたうえで、ルーブリックがデザインされてもいない。また、アメリカで開発された‘Independent WRITING Rubrics’は世界的に英語学習者全般に使用されるため、汎用的であるという利点は有する。しかしながら、この‘Independent WRITING Rubrics’は日本人で自由英作文の学習に始動期に該当する CEFR-J A2 レベルの初級英語学習者や CEFR-J B2 および C1 レベルの上級学習者の自由英作文の評価基準として必ずしも有用でないという声が聞かれる。

本研究では、CEFR-J A2 レベルの初級者と CEFR-J B2 および C1 レベルの上級者の自由英作文のルーブリックをデザインする。このルーブリックでは自由英作文の言語的特徴を重回帰分析により抽出し、それをルーブリックの「評価の観点」とする。次に、評価の観定の値に対する評価基準（到達度合い）を統計的に検証し3段階とする。上・中・下位の3段階の評価の観点に対して、重回帰分析の標準化係数 β 値（全体的評価に対する評価の観定の貢献度）の比率により、評価基準（到達度合い）の得点を設定する。このような新規性を有するルーブリックをデザインし、新たに提言するとともに、その妥当性を検証することを本研究の目的とする。

1.2.3 研究の意義

CEFR-J A2 レベルの初級者と CEFR-J B2 および C1 レベルの上級者向けのルーブリックがデザインされれば、以下に提示する項目が可能となる。

(1) CEFR-J A2 レベルの初級者のルーブリック

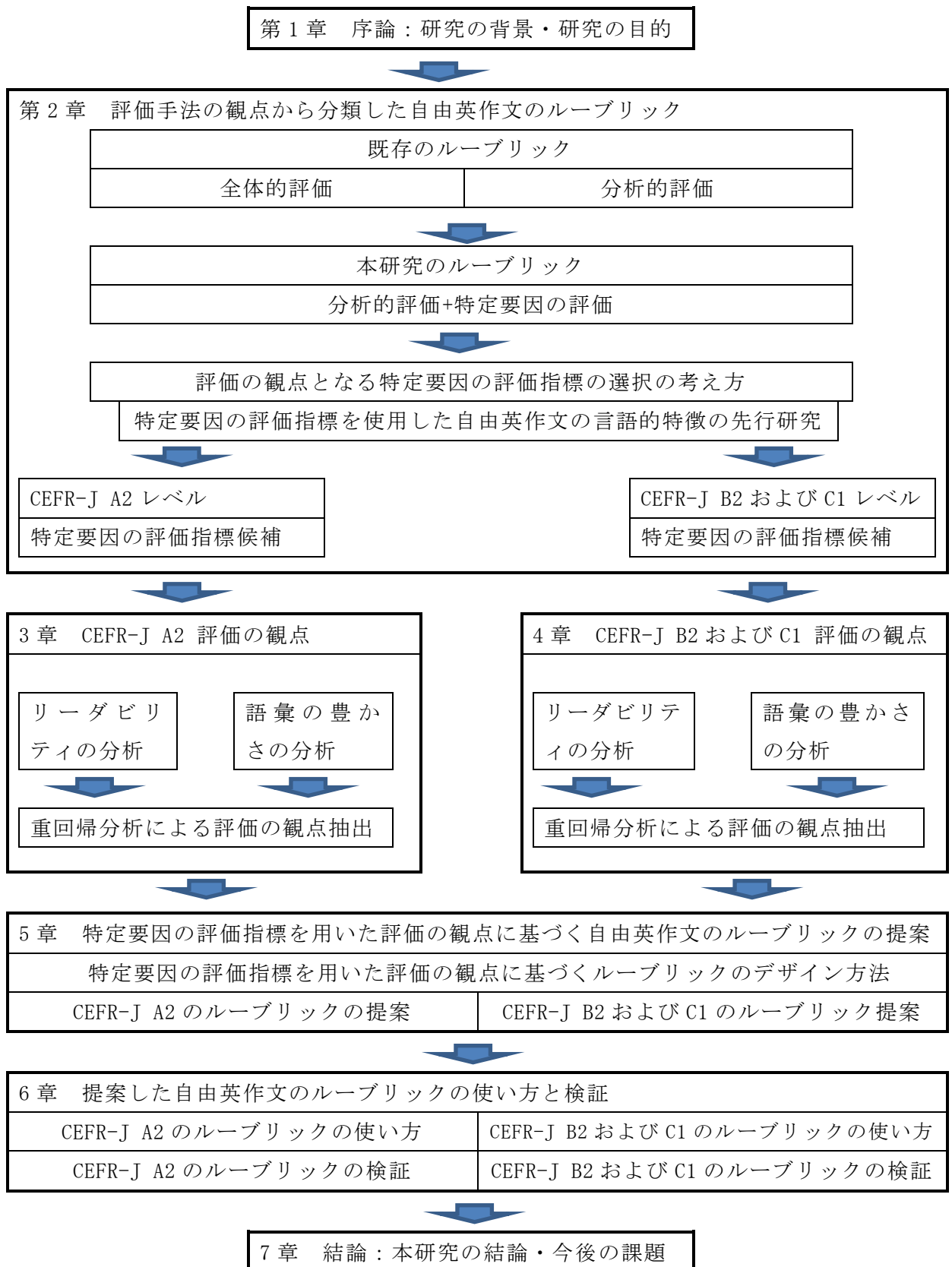
- ① ‘Independent WRITING Rubrics’では一律に評価の低い（1点）と判定された自由英作文の評価の峻別が可能となる。
- ②自由英作文の言語的特徴に基づく評価の観点は学習者に学習の指針を提供することになる。
- ③指導者による自由英作文の評価の際の効率性を高めることができる。
- ④自由英作文の言語的特徴に基づく評価の観定の定量的な数値情報は指導者に指導の指

針を提供することになる。

(2) CEFR-JB2 および C1 レベルの上級者向けのルーブリック

- ① ‘Independent WRITING Rubrics’ では高評価として 4 点ないし 5 点と判定された自由英作文をさらに峻別できる評価が可能となる。
- ②自由英作文の言語的特徴に基づく評価の観点から学習者に学習の指針を提供することになる。
- ③指導者による自由英作文の評価の際の効率性を高めることができる。
- ④自由英作文の言語的特徴に基づく評価の観点の定量的な数値情報は指導者に指導の指針を提供することになる。

1.3 本研究の全体構成



1.4 まとめ

本章では、本研究における自由英作文を定義づけし、そのメカニズムについて理論的背景を提示し、自由英作文というライティング活動が非常に複雑な要因が絡むことを示した。この研究では、日本人の初級英語学習者と上級英語学習者によって記述された自由英作文を研究の対象とするため、初級英語学習者および上級英語学習者について、欧州評議会が2001年に発行したヨーロッパにおける外国語学習、教育および評価のガイドライン The Common European Framework of Reference for Language: Learning, Teaching, Assessment (CEFR) の日本版である CEFR-J に基づき設定した。前者の初級英語学習者は CEFR-J A2 レベル（基礎段階の言語使用者の上位）、後者の上級英語学習者は CEFR-J B2（自立した言語使用者の上位）および C1（熟達した言語使用者の下位）レベルとした。このような学習者によって記述される自由英作文の重要性を提示するとともに、その普及に対する課題を示した。その最も大きな課題が複雑な要因の絡む自由英作文の評価基準の指針であるルーブリックが必ずしも十分に整備されていないことである。

本研究の目的は、上述した CEFR-J A2 レベル、CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文を対象とするルーブリックをデザインすることである。このルーブリックは、従来の自由英作文のルーブリックと異なり、自由英作文の言語的特徴を分析し、その評価の観点とし、その観点に対する到達度合いである評価基準を統計的に検証し、さらに、評価の得点は、抽出された評価の観点（言語的特徴）の標準化係数 β の比率も基に算出する。

このような自由英作文のルーブリックは学習者には学習の指針を提供し、指導者にも評価の指針を提供することになり、効率的な自由英作文の学習や指導のツールとなる可能性がある。その結果、社会的にも自由英作文の重要さは認識されているものの、その学習や指導の普及が図られていない現状を克服することになると考えられる。

2. 評価手法の観点から分類した自由英作文のルーブリック

2.1 自由英作文の3つの評価手法（全体的評価・分析的評価・特定要因の評価）

自由英作文のような一定量の英文テキストを評価する方法については全体的評価 (holistic scoring)、分析的評価 (analytical scoring)、特定要因の評価 (primary trait scoring) の3つの手法がある (Perkins, 1983)。

全体的評価は書き手によって記述された自由英作文のテキストに対する評価者の印象度に基づく評価方法である。‘Independent WRITING Rubrics’ のような評価基準に基づき、評価者の印象度によって評価がなされる。複数の採点者で評価し、評価者間の採点に関する相関係数やカップ係数などを算出し、信頼性を高める工夫が行われるのが一般的である。

分析的評価はライティングのテキストを文法、語彙、構成などの項目に分けて、評価する方法である。最も有名なものに Jacobs et al (1981) による ESL Composition Profile がある。この評価方法では、内容 (content)、構成 (organization)、語彙 (vocabulary)、言語使用 (language use)、機械的技術 (mechanics) の5つの項目に対して、30点、20点、20点、25点、5点の配点で採点がなされ、その合計点が算出される。このような分析的評価はライティング能力を構成する下位項目 (ESL Composition Profile での内容、構成、語彙、言語使用、機械的技術) ごとに評価されるため、学習者に対して診断的フィードバックを提供するという利点がある (Jacobs et al, 1981; Hamp-Lyons, 1991a)。

特定要因の評価は自由英作文を特定の要因によって評価する手法である。例えば、そのテキストが何語の単語から構成されているかという総語数が特定要因の1つとなる。1つの要因のみでテキストを評価することは問題であり、複数の要因を組み合わせることが必要となる (Hamp-Lyons, 1991 b)。この特定要因は数値を客観的に算出することができるという利点がある。

これらの3つの評価手法を使用したルーブリック (表 1-6 参照) を次項で提示し、その分析を行うこととする。

2.2 全体的評価によるルーブリック

2.2.1 馬場 (2008) のルーブリック

(1) ルーブリックの内容

馬場 (2008) では、日本人大学生の自由英作文と日本語作文の関係を分析する際に、‘Independent WRITING Rubrics’ を参考にし、日本人の英作文に見られる誤りを考慮した全体的評価を行うルーブリックを提示している。対象となったのは日本人大学生 40 名である。その英語力は測定されていないが、自由英作文のテキストの平均総語数が 66 語^{注5)} であることから、CEFR-J A2 レベルの下位者もしくは A1 レベル (TOEIC 225 点未満) の上位

注5) 本研究の対象とした CEFR-J A2 レベルの下位者の平均総語数が 76 語であるため、平均総語数が 66 語である馬場 (2008) の対象者の英語力を CEFR-J A2 レベルの下位者もしくは A1 レベルの上位者と推定した。

者である初級英語学習者と推定される。従って、この馬場（2008）の対象となった学習者は大学生とはいえ、CEFR-J A2 レベルの「書くこと」の CAN-DO ディスクリプタで提示している自由英作文を記述するレベルに到達していない英語力の学生を含む可能性がある。

表 2-1 に馬場（2008）のルーブリックを示した。このルーブリックでは、4 点を最高得点とし、0 点を最低得点とする 5 段階の尺度で自由英作文が評価される。4 点から 1 点に関しては、評価を示す記述は「文章構成」、「使用語彙」、「英文法」、「内容」の 4 項目によって構成されている。対象とする学習者が大学生であるにもかかわらず、英語力は CEFR-J A2 レベルの下位か A1 レベルと推定されるため、自由英作文の記述ができない場合も想定し、0 点となる評価を示す記述も作成されている。この評価を示す記述を基に、評価者の印象度により、自由英作文に 4 点～0 点までの得点を付与することになる。

表 2-1 馬場（2008）のルーブリック

4	<p>①主張と考えられる文、および支持文が書かれている。完璧ではないが、文章の展開もある。</p> <p>②使用語彙は限られているが、適切に使用されていて、スペリングミスもほとんどない。</p> <p>③文法上の誤りもほとんどない。平易な文であるが、伝えたい内容を的確にしており、読み手も理解できる。</p>
3	<p>①主張と考えられる文が書かれていて、その文を支持する文も 1 文から 3 文程度見られるが、その後の展開がない。</p> <p>②使用語彙が限られているものの、正しく使えており、スペリングの誤りが時々見られる程度である。</p> <p>③文法上の誤りが所々見られるが、local error か日本語の翻訳なので、伝えようとしている内容を理解することができる。</p>
2	<p>①主張と考えられる文が書かれているが、その文を支持する文がなく、その後の展開がない。</p> <p>②使用語彙が限られていて、日本語の翻訳と考えられる誤りが見られる。品詞やスペリングの誤りが見られる。</p> <p>③文法上の誤りが大半を占めるが、誤りの多くは local error かあるいは日本語の翻訳の影響によるものなので、言いたいことは何となくわかる。</p>
1	<p>①文が羅列されているが、話の展開がない。</p> <p>②使用語彙が乏しく、品詞やスペリングの誤りが見られる。</p> <p>③文法的に正しい文が全くない。あるいは 1 文から 2 文程度正しい文が含まれている。そのため、言いたいことがほとんど伝わってこない。</p>

0	①何も書けていない、あるいは単語が数語羅列されているのみで、意味をなさない。あるいは、トピック（テーマ）を書き写したのみである。
---	--

(2) 馬場（2008）のルーブリックの分析

馬場（2008）のルーブリックは初級英語学習者を対象としているため、簡潔である。評価者の視点からすると、平均総語数が 66 語程度で、初級の英語学習者によって記述された稚拙な自由英作文ならば、このルーブリックの簡潔な評価基準でも全体的評価を行うことはさほど困難ではないかと思われる。

しかしながら、以下の点に関して、問題があると考えられる。

- ①初級の英語学習者によって記述された自由英作文の言語的特徴を実証したうえで、作成されていない。
- ②評価の記述がやや一般的な表現が用いられているため、評価者間における信頼性に問題を生じる可能性がある。

2.2.2 ‘Independent WRITING Rubrics’

(1)ルーブリックの内容

TOEFL の Independent WRITING（自由英作文に相当）に対応するルーブリックである。対象は、上級者から初級者までを網羅する汎用的なルーブリックである。表 2-2 に ‘Independent WRITING Rubrics’ を示した。このルーブリックでは、5 点を最高得点とし、0 点を最低得点とする 6 段階の尺度で自由英作文が評価される。5 点から 1 点に関しては、評価を示す記述は「課題との関連性」、「構成」、「内容」、「英文法」、「語彙」の 5 項目によって構成されている。自由英作文の記述ができない場合も想定し、0 点となる評価を示す記述も作成されている。この評価を示す記述を基に、評価者の印象度により、自由英作文に 5 点～0 点までの得点を付与することになる。

表 2-2 ‘Independent WRITING Rubrics’

SCORE	TASK DESCRIPTION
5	<p>An essay at this level largely accomplished all of the following:</p> <ul style="list-style-type: none"> ■ Effectively address the topic and task ■ Is well organized and well developed, using clearly appropriate explanations, exemplifications and/or details ■ Displays unity, progression and coherence ■ Displays consistent facility in the use of language, demonstrating syntactic variety, appropriate word choice and idiomaticity, though it may have minor lexical or grammatical errors
4	<p>An essay at this level largely accomplished all of the following:</p> <ul style="list-style-type: none"> ■ Address the topic and task well, though some points may not be fully elaborated ■ Is generally well organized and well developed, using appropriate and sufficient explanations, exemplification and/or details ■ Display unity, progression and coherence, though it may contain occasional redundancy digression, or unclear connections ■ Displays facility in the use of language, demonstrating syntactic variety and range of vocabulary, though it will probably have occasional noticeable minor errors in structure, word form or use of idiomatic language that do not interfere with meaning
3	<p>An essay at this level is marked by one or more of the following:</p> <ul style="list-style-type: none"> ■ Address the topic and task using somewhat developed explanations, exemplifications and/or details ■ Displays unity, progression and coherence, though connection of ideas may be occasionally obscured ■ May demonstrate inconsistent facility in sentence formation and word choice that may result in lack of clarity and occasionally obscure meaning ■ May display accurate but limited range of syntactic structures and vocabulary
2	<p>An essay at this level may reveal one or more of the following</p>

	<p>weakness:</p> <ul style="list-style-type: none"> ■ Limited development in response to the topic and task ■ Inadequate organization or connection of ideas ■ Inappropriate or insufficient exemplifications, explanations or details to support or illustrate generalizations in response to the task ■ A noticeably inappropriate choice of words or word forms ■ An accumulation of errors in sentence structure and/or usage
1	<p>An essay at this level is seriously flawed by one or more of the following weakness:</p> <ul style="list-style-type: none"> ■ Serious disorganization or underdevelopment ■ Little or no detail, or irrelevant specifics, or questionable responsiveness to the task ■ Serious and frequent errors in sentence structure or usage
0	<p>An essay at this level merely copies words from the topic, rejects the topic, or is otherwise not connected to the topic, is written in a foreign language, consists of keystroke, characters, or is blank.</p>

(2) ‘Independent WRITING Rubrics’ の分析

このルーブリックは TOEFL の Independent WRITING 型課題のために開発されたものであり、世界中で使用されているとおり、初級英語学習者から上級英語学習者までを網羅した汎用性に優れたルーブリックと言える。

しかしながら、以下の点に関して、問題があると考える。

- ①Nishijima, et al (2007)によって、評価の記述の表現があまりに一般化され過ぎているという批判があるように、この記述から印象度による全体的評価を行う場合、評価者間の信頼性に支障をきたす可能性がある。例えば、高評価の範疇に入る自由英作文の評価で、5点と4点に関して、評価者間の印象度が異なる危険性がある。
- ②評価の記述の表現があまりに一般化され過ぎているために、自由英作文の学習に対する十分な指針を学習者に提供できない可能性がある。
- ③評価の記述の表現があまりに一般化され過ぎているために、自由英作文の指導に対する十分な指針を指導者に提供できない可能性がある。
- ④英語学習者によって記述された自由英作文の言語的特徴を実証したうえで、作成されていない。
- ⑤CEFR-J A2 レベルのような初級の英語力の学習者によって記述された自由英作文は大

半が1点となり、実際には英文の質に差のある場合でも、評価の峻別がなされないことが懸念される。

2.3 分析的評価によるルーブリック

2.3.1 ESL Composition Profile

(1) ルーブリックの内容

ESL Composition Profile は Jacobs et al (1981) によって作成されたルーブリックである。分析的評価を採用している最もよく知られたルーブリックである。ESL (English as a Second Language) ^{注6)} ということから第2言語として英語を学ぶ学習者を対象として作成されている。表 2-3 に ESL Composition Profile を示した。このルーブリックでは、内容 (content)、構成 (organization)、語彙 (vocabulary)、言語使用 (language use)、機械的技術 (mechanics) の5つの項目に対して、それぞれ、30点、20点、20点、25点、5点の配点で採点が行われ、その合計点が算出され、最高点は100点、最低点は34点となる。5つの項目に関して、‘EXCELLENT TO VERY GOOD’、‘GOOD TO AVERAGE’、‘FAIR TO POOR’、‘VERY POOR’の4段階に分かれており、配点が決めている。内容の30点は4段階(30点-27点、26点-22点、21点-17点、16点-13点)、構成の20点は4段階(20点-18点、17点-14点、13点-10点、9点-7点)、語彙の20点は4段階(20点-18点、17点-14点、13点-10点、9点-7点)、言語使用の25点は4段階(25点-22点、21点-18点、17点-11点、10点-5点)、機械的技術の5点は4段階(5点、4点、3点、2点)の各配点で、得点が付与されることになる。

表 2-3 ESL Composition Profile

	LEVEL	CRITERIA
CONTENT	30-27	EXCELLENT TO VERY GOOD: knowledgeable • substantive • thorough • development of thesis • relevant to assigned topic
	26-22	GOOD TO AVERAGE: some knowledge of subject • adequate range • limited development of thesis • mostly relevant to topic, but lacks detail
	21-17	FAIR TO POOR: limited knowledge of subject • little substance • inadequate development of topic
	16-13	VERY POOR: does not show knowledge of subject •

^{注6)} 米山 (2003) によると、ESL (English as a Second Language) は「英語が学校の授業の媒体として用いられ、政治、経済、法律などの分野で公用語として使用されており、さらに日常生活の場面でも使われている地域で英語が学習されている状況を指す」と定義されている。

		non-substantive • not pertinent • OR not enough to evaluate
ORGANIZATION	20-18	EXCELLENT TO VERY GOOD: fluent expression • ideas clearly stated/supported • succinct • well-organized • logical sequencing • cohesive
	17-14	GOOD TO AVERAGE: somewhat choppy • loosely organized but main ideas stand out • limited support • logical but incomplete sequencing
	13-10	FAIR TO POOR: non-fluent • ideas confused or disconnected • lacks logical sequencing and development
	9-7	VERY POOR: does not communicate • no organization • OR not enough to evaluate
VOCABULARY	20-18	EXCELLENT TO VERY GOOD: sophisticated range • effective word/idiom choice and usage • word form mastery • appropriate register
	17-14	GOOD TO AVERAGE: adequate range • occasional errors of word/idiom form, choice, usage but meaning not obscured
	13-10	FAIR TO POOR: limited range • frequent errors of word/idiom form, choice, usage • meaning confused or obscured
	9-7	VERY POOR: essentially translation • little knowledge of English vocabulary, idioms, word form • OR not enough to evaluate
LANGUAGE USE	25-22	EXCELLENT TO VERY GOOD: effective complex constructions • few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions
	21-18	GOOD TO AVERAGE: effective but simple constructions • minor problems in complex constructions • several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions but meaning seldom obscured
	17-11	FAIR TO POOR: major problems in simple/complex

		constructions • frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions, and/or fragments, run-ons, deletions • meaning confused or obscured
	10-5	VERY POOR: virtually no mastery of sentence construction rules • dominated by errors • does not communicate • OR not enough to evaluate
MECHANICS	5	EXCELLENT TO VERY GOOD: demonstrates mastery of conventions • few errors of spelling, punctuation, capitalization, paraphrasing
	4	GOOD TO AVERAGE: occasional errors of spelling, punctuation, capitalization, paraphrasing but meaning not obscured
	3	FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, paraphrasing • poor handwriting • meaning confused or obscured
	2	VERY POOR: no mastery of conventions • dominated by errors of spelling, punctuation, capitalization, paraphrasing • handwriting illegible • OR not enough to evaluate

(2) ESL Composition Profile の分析

この ESL Composition Profile は、分析的評価を使用した最もよく知られたルーブリックであり、全体的評価の ‘Independent WRITING Rubrics’ とともに、初級英語学習者から上級英語学習者までを網羅した汎用性に優れたルーブリックと言える。分析的評価のため、下位項目（ESL Composition Profile での内容、構成、語彙、言語使用、機械的技術）ごとに評価されるため、学習者に対して診断的フィードバックを提供するという利点がある (Jacobs et al, 1981; Hamp-Lyons, 1991a)。

しかしながら、以下の点に関して、問題があると考えられる。

- ① 評価者が 5 つの下位項目の 4 段階のクライテリア (‘EXCELLENT TO VERY GOOD’ 、 ‘GOOD TO AVERAGE’ 、 ‘FAIR TO POOR’ 、 ‘VERY POOR’) の判断に戸惑う可能性がある。
- ② 評価者が 5 つの下位項目の 4 段階のクライテリア内で何点を付与すべきかについても戸惑う可能性がある。例えば、評価者が対象とする自由英作文における内容の項目で ‘EXCELLENT TO VERY GOOD’ と判断した場合、30 点-27 点の得点を付与することにな

るが、その得点の判断の基準が示されていない。

- ③英語学習者によって記述された自由英作文の言語的特徴を実証したうえで、作成されていない。

2.3.2 Nakanishi (2006)のルーブリック

(1)ルーブリックの内容

Nakanishi (2006)では、ESL Composition Profile を改良した EFL Composition Profile を作成している。これは、第2言語 (English as a second language: ESL) ではなく外国語 (English as a foreign language: EFL^{注7)}) として英語を学ぶ日本人大学生を意識して作成されたルーブリックである。対象となった学習者の英語力は測定されていない。表2-4にEFL Composition Profileを示した。このルーブリックでは、ESL Composition Profile に準拠して、内容 (content)、構成 (organization)、語彙 (vocabulary)、言語使用 (language use)、機械的技術 (mechanics) の5つの項目に対して、それぞれ、30点、20点、20点、25点、5点の配点で採点がなされ、その合計点が算出され、最高点は100点、最低点は24点となる。

表 2-4 EFL Composition Profile

5 points: Excellent	4 points: Very good	3 points: Average	2 points: Fair	1 point: Poor
Criteria	Checklist			Score: Weighting
Content	Does a writer answer a writing topic appropriately? Is all the necessary information (who, what, where, how) included? Do you think the foreign student who reads the description will be able to play "Janken" ?			(Points)×6
Organization	Does a draft have a topic sentence? Does a draft develop based on the topic sentence? Are logical connectors used appropriately? Is the order of a draft appropriate?			(Points)×4
Vocabulary	Are the choices of words appropriate?			(Points)×4

注7) 米山 (2003) によると、EFL (English as a Foreign Language) は「英語が学校の教科として教えられているが、授業の媒体とならず、また、政治、経済、法律などの分野で公用語として用いられることがなく、通常の日常生活のコミュニケーションのための言語ともなっていない地域で英語が学習される状況を指す」と定義されている。

	Are a variety of words and expression used?	
Language Use	Is grammar (subject-verb agreement, tense, numeral, article, preposition, pronoun, and reflexive pronoun) accurate?	(Points)×5
Mechanics	Are paragraphing, punctuation, spelling, and capitalization accurate?	(Points)×1

5 criteria and the weighting of each were based on ESL Composition Profile.

注：Contentに‘Do you think the foreign students who reads the description will be able to play “Janken”?’とあるのは、このトピックが「ジャンケンを外国人にどのように説明するか」であるためである。

(2) EFL Composition Profile の分析

ESL Composition Profile と同様に、分析的評価のため、下位項目ごとに評価されるため、学習者に対して診断的フィードバックを提供するという利点がある (Jacobs et al, 1981; Hamp-Lyons, 1991a)。また、各項目に対して、5段階 (Excellent:5点、Very good:4点、Average:3点、Fair:2点、Poor:1点) で得点を付与し、その後、各項目に重み付けをした数値 (内容:6倍、構成:4倍、語彙:4倍、言語使用:5倍、機械的技術:1倍) を掛け合わせるという算出方式により、「2.3.1 ESL Composition Profile」の「(2) ESL Composition Profile の分析」で示した問題点の②を改善することが可能となっている。さらに、評価の記述も ESL Composition Profile よりも単純になっており、評価者にとっては評価がしやすくなっている。

しかしながら、以下の点に関して、問題があると考える。

- ①ESL Composition Profile よりも単純化された評価を基に5段階の評価を加えなければならぬため、評価者間による評価の信頼性を欠く可能性がある。
- ②英語学習者によって記述された自由英作文の言語的特徴を実証したうえで、作成されていない。

2.3.3 Pilot Rubric for Paragraph / Essay Writing

(1) ルーブリックの内容

Pilot Rubric for Paragraph / Essay Writing は Nishijima, et al (2007) で作成されたルーブリックである。このルーブリックはアメリカの2つのパブリックスクールの第1言語向けのルーブリックを結合した後に、日本人の大学生の自由英作文の評価基準となるように修正を加え、Pilot Rubric for Paragraph/Essay Writing として提示している。分析的評価のため、下位項目ごとに評価されるため、学習者に対して診断的フィードバックを提供するという利点がある (Jacobs et al, 1981; Hamp-Lyons, 1991a)。彼らはこれを実

際に使用していく中で、今後の研究を進めていくとしている。表 2-5 に Pilot Rubric for Paragraph / Essay Writing を示した。このルーブリックでは、分析的評価の下位項目として、内容・展開 (Content and Idea Development)、構成 (Organization)、文法 (Grammar)、語彙 (Vocabulary)、綴り・句読点など (Mechanics) の 5 つの項目に対して、それぞれ 4 段階の尺度となっている。それぞれの項目の最上位が 4 点、最下位が 1 になっているため、最高点は 20 点、最低点は 5 点となる。

表 2-5 Pilot Rubric for Paragraph / Essay Writing

dimension/domain	1	2	3	4
Content and Idea Development 内容・展開	メッセージが不明確である。またテーマについての知識もほとんど伺えない。理由付けは不十分で、テーマと直接関係ないものや矛盾したものが多数見られる。	メッセージは十分伝わっていない。またテーマについての知識も限定されたものである。理由付けは不十分で、予測可能ものが多数見られる。	メッセージが明確に伝わっている。しかし、テーマについて十分な知識を持ち合わせているとは言えない。理由付けは適切であるが、一般的すぎるものやテーマとは直接関係のないものを含んでいる。	メッセージは明確に伝わっている。テーマについて十分な知識を持ち合わせている。理由付けは十分あり、説得力ある情報を使っている。
Organization 構成	文章の構成が全くみられない。	文章の構成は幾分不明確で論理性に欠ける。	文章の構成は全体的に明確で論理的である。	文章の構成 (序論、本論、結論) は明確で論理的である。
Grammar 文法	(1) 文と文の関連性が不明確である。 (2) 反復的な文が極めて多く使われている。	(1) 文と文の関連性がいくぶん不明確である。 (2) 反復の多い文が使われて	(1) 文と文の関連性がほぼ明確に示されている。 (2) 多様な文が使われている。	(1) 文と文の関連性が明確に示されている。 (2) 多様な文が適切に使われ

	(3) 文構造は正しくない。 (4) 全体にわたって誤りが多い。	いる。 (3) 文構造があまり正しくない。 (4) 主語・動詞の一致、時制、代名詞、冠詞、前置詞などにおいて頻繁に誤りがある。	(3) 文構造はほぼ正しい。 (4) 主語・動詞の一致、時制、代名詞、冠詞、前置詞などにおいていくつかの誤りがある。	ている。 (3) 文構造は正しい。 (4) 主語・動詞の一致、時制、代名詞、冠詞、前置詞など誤りがない。
Vocabulary 語彙	語彙や慣用表現の範囲は極めてかぎられている。語の用法や語形の不適切さが目立つ。	語彙や慣用表現の範囲は狭い。語や慣用表現の形、選択、および用法において誤りが頻繁に見られる。	ある程度広範囲の語彙や慣用表現が使用されている。語や慣用表現の形、選択、および用法において時折誤りが見られる。	語彙や慣用表現が豊富で適切に使用されている。
Mechanics 綴り、句読点など	綴り、句読点、大文字の使用、段落分けが不正確である。	綴り、句読点、大文字の使用、段落分けがいくぶん不正確である。	綴り、句読点、大文字の使用、段落分けがだいたい正確である。	綴り、句読点、大文字の使用、段落分けが正確である。

(2) Pilot Rubric for Paragraph / Essay Writing の分析

Nishijima, et al (2007)は‘Independent WRITING Rubrics’の評価の記述の表現が一般的かつ抽象的ななっていることを批判しており、彼らの作成した Pilot Rubric for Paragraph / Essay Writing というルーブリックでは評価の記述が、その点も配慮して、より具体的な表現が使われている。アメリカの2つのパブリックスクールにおける第1言語のルーブリックを結合した後に、日本人の大学生の自由英作文の評価基準となるように修正を加えているため、CEFR-J B1 レベルの中級者向けのルーブリックとしては有用ではないかと考えられる。また、ESL Composition Profile や EFL Composition Profile と同様に、分析的評価のため、下位項目ごとに評価されるため、学習者に対して診断的フィードバックを提供するという利点がある(Jacobs et al, 1981; Hamp-Lyons, 1991a)。

しかしながら、以下の点に関して、問題があると考える。

- ①分析的評価の下位項目の内容・展開、構成、文法、語彙、綴り・句読点と同じ重みで得点が算出されていることに疑問がある。例えば、ESL Composition Profile や EFL Composition Profile における機械的技術は他の下位項目よりも重みが低くなっているが、Pilot Rubric for Paragraph / Essay Writing の機械的技術に相当する綴り・句読点が他の下位項目と同じ重みになっていることである。
- ②英語学習者によって記述された自由英作文の言語的特徴を実証したうえで、作成されていない。

2.4 分析的評価での評価の観点に特定要因の評価指標を用いたルーブリックの考え方

本研究におけるルーブリックは自由英作文の評価手法においては特定要因の評価手法に基づく分析的評価を採用する。自由英作文の言語的特徴を重回帰分析によって抽出した説明変数を特定要因の評価指標とし、それをルーブリックの評価の観点とし、分析的評価の各項目とする。このことは以下の利点がある。

- (1) 分析的評価を採用したことにより、下位項目ごとに評価されるため、学習者に対して診断的フィードバックを提供するという利点がある。(Jacobs et al, 1981; Hamp-Lyons, 1991a)。
- (2) 評価の観点が学習者によって記述された自由英作文の言語的特徴に基づくことになる。さらに、本研究では、既存のルーブリックでは全く扱われていなかった評価基準（到達の度合い）に関する実証された数値に基づいて、その水準を定めることとした。このことによって以下の利点がある。
 - ①評価の観点の数値化により評価基準（到達の度合い）の水準を定めるので、記述の説明のみの場合に比較して、その自由英作文がどの水準に入るのかが正確に把握することができる。
 - ②評価の観点の数値化により評価基準（到達の度合い）を設定できるので、信頼性と客観性が高まる。
 - ③評価の観点の数値化は、簡潔な記述を求められるルーブリックの評価の記述を補完することになる。

このようなルーブリックをデザインするために、以下の分析が必要になる。

- (1) 評価の観点を作成するために、自由英作文の言語的特徴を分析すること。
- (2) 評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値すること。
- (3) 評価の観点に対する評価基準（到達度合い）の得点を重回帰分析の標準化係数 β 値（全体的評価に対する評価の観点の貢献度）の比率により設定すること。

この(1)については、本章 5 節で扱い、(2)および(3)については、3 章および 4 章で扱うこととする。

2.5 評価の観点となる特定要因の評価指標候補の選択の考え方

2.5.1 自由英作文の言語的特徴に関する先行研究

本研究では、ルーブリックの評価の観点は、対象とする自由英作文の言語的特徴を実証することによって抽出された特定要因の評価指標を採用することとする。そこで、まず、自由英作文の言語的特徴を表す特定要因の評価指標を選択するにあたり、自由英作文の言語的特徴に関する内外の代表的な先行研究を以下に概観する。そのような先行研究を基に、本研究の対象とする CEFR-J A2 レベルと CEFR-J B2 および C1 レベルの学習者によって記述される自由英作文の言語的特徴を表す特定要因の評価指標候補を選択する。

海外の研究では Ferris(1994)があり、英語力の異なるレベルの書き手 160 人（アラビア語、中国語、日本語、スペイン語を母語とする英語を外国語とする学習者で海外に留学している学生のため、英語力は少なくとも CEFR-J B2 レベル以上と考えられる）による英作文を対象として、28 個の言語的特徴に関する変数を説明変数候補とし、その全体的評価を従属変数とする重回帰分析を行っている。結果として、5 つの説明変数（総語数、同義語・反意語、語長、受動態、3 人称・非人称代名詞）により $R^2=0.50$ のモデルが構築された。その説明変数で圧倒的であったのは総語数であった（ R^2 に 38%程度貢献）。

日本では、石田(2013)が英語の 125 人の非母語話者と 125 人の母語話者によって記述された英作文を対象とし、判別分析により両者を比較検証している。その結果、総語数、異なり語数、Guiraud Index（語彙の多様性を表す指標）と 2 語表現の総数が両者を判別する言語的特徴であるとしている。また、杉浦（2008）では、日本語を母語とする大学生および大学院生 48 人によって記述された英文エッセイを対象として重回帰分析を使って、言語的特徴を調査している。結果として、総語数、文あたりの従属節数、文あたりの接続語句数、平均文長が、その言語的特徴を示すとしている。しかし、重回帰分析の自由度修正済み決定係数が 0.44 であり、回帰式の当てはまり度はそれほど高くはないとしている。これらの研究対象者の英語力は CEFR-J では、英語の非母語話者は少なくとも B1 レベル、大半は B2 レベル、英語の母語話者は C1 レベルないし C2 レベルと考えられ、英語の非母語話者であってもかなり高度な英語力を有すると考えられる。これに対して、水本（2008）は英語力 300 点前半の英語力（リーディングおよびリスニング各問題 25 問の 4 択問題から構成される TOEIC に準ずるテストにより測定）の大学生 40 人に自由英作文を書かせている。その英語力は CEFR-J A2 レベルとは考えられるが、A2 レベルの英語力（TOEIC 225 点～550 点）全般をカバーしているとは言い難い。この研究では、自由英作文の全体的評価を従属変数、総語数、Type-Token ratio (TTR: 語彙の多様性を表す指標)、標準化 TTR、平均語長、D（語彙の多様性を表す指標）、Flesch-Kincaid Grade (FKG: リーダビリティの指標)、J8 Index（『JACET 8000^{注8)}』における難易語の使用割合）の 7 つを独立変数候補として重回帰分析を行っている。結果として、総語数、FKG、平均語長の 3 変数が説明変数として抽出され、それぞれの標準化係数 β 値は 0.68、0.44、-0.31 で、調整済み $R^2=0.60$ であるとしている。

Ferris(1994)、杉浦（2008）、水本（2008）は対象となる学習者の英語力にはかなりの

注8) 大学英語教育学会（JACET）によって編纂された語彙集。日本人が英語でコミュニケーションを図るために必要な 8000 語を選定している。

相違があるにも関わらず、言語的特徴としての総語数が共通して自由英作文の説明変数となっている。石田(2013)では非母語話者と母語話者によって記述された自由英作文の比較を研究課題としているが、その場合においても総語数が、非母語話者と母語話者を判別する変数として、挙げられている。このことは、自由英作文において情報量としての総語数が、基礎となることを示していると考えられる。表 2-6 にこれらの先行研究の結果をまとめた。

表 2-6 自由英作文の言語的特徴に関する先行研究

研究	対象者（母語）	英語力	分析手法	結果
Ferris(1994)	アメリカの大学へ留学している英語学習者 160 人（アラビア語、中国語、日本語、スペイン語）	CEFR-J B2 レベル以上	重回帰分析 $R^2=0.50$	総語数、同義語・反意語、語長要因、受動態、3 人称・非人称代名詞が自由英作文の言語的特徴
石田（2013）	母語話者 125 人（英語） 非母語話者 125 人（日本語）	母語話者は CEFR-J C1 以上、 非母語話者は CEFR-J B1 以上	判別分析 正変別率 96%	総語数、異なり語数、Guiraud Index、2 語表現の総数が両者を判別する言語的特徴
杉浦（2008）	日本人の大学生および大学院生 48 人（日本語）	CEFR-J B1 以上	重回帰分析 自由度修正済み 決定係数：0.44	総語数、文あたりの従属節数、文あたりの接続語句数、平均文長が自由英作文の言語的特徴
水本（2008）	日本人の大学生 40 人（日本語）	CEFR-J A2 下位レベル	重回帰分析 調整済み $R^2=0.60$	総語数、FKG、平均語長が自由英作文の言語的特徴

本研究においても、これらの先行研究で自由英作文の言語的特徴を抽出するために使われた重回帰分析を利用することとする。その重回帰分析では自由英作文の全体的評価を目的変数とし、決定係数 R^2 が最低限 0.50 を超えることを目処とし、説明変数候補（特定要因の評価指標）を選択することとする。

2.5.2 CEFR-J A2 レベル自由英作文の評価の観点となる特定要因の評価指標候補

CEFR-J A2 レベルの自由英作文の言語的特徴に関する説明変数候補の選択に関しては、「2.5.1 自由英作文の言語的特徴に関する先行研究」で使用されている特定要因の評価指標に加えて、自由英作文に関する別の視点からの研究で使われている特定要因の評価指標を参考にした。その結果、Fluency、文、談話、語彙、リーダビリティの5つの視点から15変数を説明変数候補として選択することとした。

- (1) Fluency の視点からは、総語数および総文数を説明変数候補とした。
- (2) 文の視点からは、1文の平均語数、Error-free T-unit 平均語数を説明変数候補とした。
- (3) 談話 (discourse^{注9)}) の視点からは、Halliday and Hasan (1976)における結束性に基づき、1英文テキストにおける結束数の合計、結束数の合計に占める指示 (R) および接続 (C) の割合を説明変数候補とした。
- (4) リーダビリティの視点からはリーダビリティを表す指標である Flesch-Kincaid Grade (FKG)、Automated Readability Index (ARI)、Coleman-Liau Index (CLI) を説明変数候補とした。
- (5) 語彙の視点からは lexical variation (語彙の多様性) を表す Guiraud Index (以下、GI とする)、D、the Measure of Textual Lexical Diversity (以下、MTLD とする)、lexical sophistication (語彙の広範さ) を表す P_Lex および S、平均語長を説明変数候補とした。

以下に自由英作文の言語的特徴に関する説明変数候補の選択についての5つの視点について詳述する。

- (1) の Fluency (流暢さというより、書かれた情報量) の視点からは、3.1.1 の自由英作文の言語的特徴のいずれの研究においても説明変数となっていた総語数を採用した。総文数という特定要因の評価指標も Spaan (1993)、Homburg (1984)、Ishikawa (1995) などの研究で使用されており、説明変数候補とした。総文数という名称からすると、(2) の文の視点とも考えられなくはないが、書かれた情報量としての総文数と考え、Fluency の視点の範疇に入れることとした。
- (2) の文の視点からの1文の平均語数は Ferris (1994)、Reid (1990)、Song & Caruso (1996) などの研究でも使われており、説明変数候補として採用した。Error-free T-unit 平均語数はライティングの研究において非常によく使用される特定要因の評価指標 (武田, 1999; Kanatani. et al., 1993; Homburg, 1984; Ishikawa, 1995 など) である。T-unit^{注10)} の有効性には賛否があるが、本研究では「エラー^{注11)} のない T-unit の平均の長さ

注9) 米山 (2003) によれば、コミュニケーションの目的を持った、文を超えたまとまりのある言語単位。

注10) T-unit とは、Hunt (1965) によって考案され、1つの独立節 (S+V の構造) とすれに付加されるあらゆる種類の従属節 (独立節に含まれる S+V の構造) からなる、文を分割す

総数が増加しておれば、かなり自信をもって作文力が向上していると生徒に知らせることができる」(沖原, 1994, p. 31) という主張を考慮し、Error-free T-unit 平均語数を採用する。

(3) の談話の視点から、結束性に関する特定要因の評価指標を選択したのは、橋内 (1995: 15) が英語を外国語とする初級学習者の自由英作文においては「まずは cohesion(結束性)について十分理解することが重要である」と指摘おり、これは CEFR-J A2 レベルのような初級学習者によって記述される自由英作文においては結束性^{注12)}が重要になるためである。結束性に関しては、先行研究では自由英作文の特定要因の評価指標として使われたものがないため、Halliday & Hasan (1976)を基に、指標を作成することとする。Halliday & Hasan (1976)では英語における結束性を指示(reference: R)、代用(substitution: S)、省略(ellipsis: E)、接続(conjunction: C)、語彙的結束性(lexical cohesion)の5つを挙げている。この5つの結束性の種類から、結束数の合計^{注13)}、結束数の合計に占める指示(R)および接続(C)の割合^{注14)}という2つの自由英作文の特定要因の評価指標を考案する。

(4) のリーダビリティ^{注15)}に関しては、FKG、ARI、CLIの各指標では、対象とするテキストを読むために最低限必要とされる教育レベルをアメリカの学年によって示される(染

る最小単位である(Richards et al, 1992, p. 390)。①と②の文は同じ意味を表すが、①はsoが等位接続詞であるため、so以後は新たな独立節となり、2 T-unitで11語のため、1 T-unitあたり5.5語となる。一方、②では、becauseは従属接続詞のため、because以下も従属節となり、主となる独立節の中に含まれ、1 T-unitで11語となる。T-unit内の単語が多ければ、それだけ熟達した作文と見なされる。従って、②の方が熟達した英文となる(相沢, 2001, p. 68)。①I had a cold yesterday, so I couldn't go to school. ②I couldn't go to school yesterday because I had a cold. Error-free T-unit 平均語数は自由英作文の中で、エラーのないT-unitの平均語数となる。

^{注11)} CEFR-J A2 レベルの初級の学習者であるため、「文もしくは発話の理解を難しくするか、もしくは全く理解できなくするような文構造の主要素の使い方の誤り」(Richards et al., 1992)というグローバルエラーを意味することとする。

^{注12)} この定義として「既出あるいは既知のものとされている情報の全部ないし一部が後出の部分に受け継がれるか、前後の言語表現の意味は意味的に許容される何らかの関係によって結びつけることによって文の集合をテキストとして成り立たせている意味関係」(中島・大塚, 1982: 197)とした。

^{注13)} 結束数の算出の具体例を挙げると、“I've been to America to see my aunt. She is getting very forgetful these days.”(調査参加者以外のA高等学校の生徒によって記述された英文)という2つの英文では、my auntとSheが指示的結束性となっており、「第1文と第2文には指示的結束が1つある」として結束数を算出した。

^{注14)} 本研究の対象としたCEFR-J A2 レベルの学習者によって記述された自由英作文のテキストでは指示(R)と接続(C)に結束数が全結束数の70%程度を占めるので、その割合は自由英作文の言語的特徴になるのではないかと推定した。

^{注15)} 「テキストの読みやすさ、理解しやすさ」(小池, 2003)であり、「語彙の難易、文長等の文体による読みやすさ」(高梨・卯城, 2000)である。

谷, 2009)。従って、その数値が例えば、8.0 の場合には、そのテキストを読むためにはアメリカの学校制度における 8 年生程度の教育レベルが必要ということになる。FKG は音節数に依拠した公式によって算出され、ARI と CLI は音節数に依拠しない音節によって算出される。これらのリーダビリティに関する指標については 3 章 1 節において詳述する。

(5) の語彙の視点に関しては、多様性^{注16)}と広範さ^{注17)}という語彙の豊かさの表す指標を選択した。前者の語彙の多様性に関しては GI、D、MTLD、後者の語彙の広範さに関する指標としては P_Lex および S を採用した。これらの語彙の多様性および広範さを表す指標については 3 章 2 節において詳述する。また、平均語長は水本 (2008) において、言語的特徴の説明変数として抽出されているため、本研究においても説明変数候補とした。このようにして選択した特定要因の評価指標を説明変数候補として、自由英作文の全体的評価を目的変数とする重回帰分析を行い、CEFR-J A2 レベルの自由英作文の言語的特徴を分析する。この 5 つの視点からの説明変数候補であれば、重回帰分析の結果、決定係数 R^2 が最低限 0.50 を超えることは可能であると考えられる。

その重回帰分析の前段階として、CEFR-J A2 レベルの自由英作文をリーダビリティと語彙の豊かさ (多様性と広範さ) の視点からもより詳細な分析を試みることにする。

前者のリーダビリティの分析は、CEFR-J A2 レベルの自由英作文の言語的特徴を表す説明変数候補の特定要因の評価指標として、FKG、ARI、CLI の 3 つの指標でどれが適しているかの判断にも役立つと考えられる。また、日本では、リーダビリティの視点から自由英作文の分析した研究は水本 (2008) の他には樋口 (2010) しかなく、新たな知見が得られれば、本研究の主目的である CEFR-J A2 レベルの自由英作文のルーブリックのデザインの基礎データになりうるとも考えられる。

後者の語彙の豊かさ (多様性と広範さ) の視点から分析は、CEFR-J A2 レベルの自由英作文の言語的特徴を表す説明変数候補の特定要因の評価指標として、語彙の多様性を表す指標である GI、D、MTLD ではどれが、また、語彙の広範さを表す指標である P_Lex と S ではどちらが、適しているのかの判断にも役立つと考えられる。また、日本では、語彙の広範さの視点から自由英作文の分析した研究は小島 (2013) しかなく、新たな知見が得られれば、本研究の主目的である CEFR-J A2 レベルの自由英作文のルーブリックのデザインの基礎データになりうるとも考えられる。

注16) この指標では、テキストにおける異語数と総語数に基づき、どの程度多様な語が使用されているかを評価することになる (小島, 2013)。

注17) この指標では、語彙の一般的な頻度情報を利用して、頻度の観点から語彙使用がどの程度広範なものかを評価することになる (小島, 2013)。

2.5.3 CEFR-J B2 および C1 レベルの自由英作文の評価の観点となる特定要因の評価指標候補

CEFR-J B2 および C1 レベルの自由英作文の言語的特徴に関する説明変数候補の選択に関しては、「2.5.1 の自由英作文の言語的特徴に関する先行研究」で使用されている特定要因の評価指標に加えて、自由英作文に関する別の視点からの研究で使われている特定要因の評価指標を参考にした。

自由英作文の言語的特徴に関する変数の選択に関しては、上級英語学習者であることを考慮した。上級英語学習者の自由英作文の特徴は、Educational Testing Service（以下、ETS）によって作成された‘Independent WRITING Rubrics’の5点のTASK DESCRIPTIONに示されていると考えた。それを要約すると、「構成が整っており、適切な説明や例証がなされている」、「内容が首尾一貫している」、「言語使用に関して、語彙の選択も的確で語彙や文法の誤りも少ない」とされている。これらの記述を出来る限り、包括できるように、言語的特徴に関する変数を抽出するようにした。そのために構成、内容、言語使用、Fluency、文、語彙、リーダビリティの7つの視点から、15変数を選択することとした。

- (1) 構成の視点から、上級英語学習者によって記述された自由英作文を考慮して、その「構成」を表す特定要因の評価指標を独自に作成した。以下にその作成方法を記述する。序論と結論の体裁が整っていることを前提とする。序論については、与えられたトピックを踏まえ、本論につながる構成をしているかどうかを2点満点（0.5点刻み）で採点し、結論については、本論と首尾一貫したつながりを有しているかどうかを2点満点（0.5点刻み）で採点する。従って、「構成」は最高点4点、最低点0点となる特定要因の評価指標となる。
- (2) 内容の視点から、上級英語学習者によって記述された自由英作文を考慮して、その「内容」を表す特定要因の評価指標を独自に作成した。以下にその作成方法を記述する。本論を主題部分と支持部分に分け、主題部分の内容が適切かどうかを3点満点（0.5点刻み）で採点し、支持部分の内容が適切かどうかを3点満点（0.5点刻み）で採点する。従って、「内容」は最高点6点、最低点0点となる特定要因の評価指標となる。
- (3) 言語使用の視点から、上級英語学習者によって記述された自由英作文を考慮して、その「言語使用」を表す特定要因の評価指標を独自に作成した。以下にその作成方法を記述する。語彙や文法の誤りに関して、エラーの数で5段階の得点を付与した。平均総語数が約230語、1文平均語数が約12語（表4-11参照）のため、5点は「対象とする自由英作文のテキスト内40語（3文程度）以上で平均1つのエラー」、4点は「対象とする自由英作文のテキスト内20語～39語で平均1つのエラー」、3点は「対象とする自由英作文のテキスト内15語～19語で平均1つのエラー」、2点は「対象とする自由英作文のテキスト内10語～14語で平均1つのエラー」、1点は「対象とする自由英作文のテキスト内10語未満で平均1つのエラー」とした。これはエラーの数により、評価者の印象度を考慮したものである。この方式を採用すると総語数の増減によっても対処でき、

汎用性があると考えらる。

- (4) Fluency の視点からは総語数および文数を説明変数候補とした。
- (5) 文の視点からは 1 文の平均語数を説明変数候補とした。
- (6) 語彙の視点からは lexical variation (語彙の多様性) を表す GI、D、および MTL、lexical sophistication (語彙の広範さ) を表す P_Lex および S、平均語長を説明変数候補とした。
- (7) リーダビリティの視点から FKG、ARI、CLI を説明変数候補とした。

以下に自由英作文の言語的特徴に関する説明変数候補の選択についての 7 つの視点について詳述する。(1)～(3)の「構成」、「内容」、「言語使用」については、次のように考えた。「構成」に関しては、序論、本論、結論の形式で記述されている自由英作文を前提とした。「構成」という視点からすると序論および結論という体裁を成していれば、本論も必然的に存在すると考え、序論および結論の部分を採点の対象とする変数を作成する。「内容」に関しては、本論の部分を対象とし、それを主題部分と支持部分に分け、採点の対象とする変数を作成する。「言語使用」に関しては、本研究の対象となる CEFR-J B2 および C1 レベルの学習者は TOEIC の平均得点で 907 点 (表 4-2 参照) であるため、語彙や文法のローカルエラーまでも誤りとする厳格なものとし、エラーの重み付けはせず、誤りがあればエラーとして算出する。これらの変数に関する配点については、Jacob et al (1981) の ESL Composition Profile の内容、構成、言語使用に関して、30 点、20 点、25 点の各配点比率は一般化されているものと考え、その比率に準拠させることとし、「内容」は 6 点満点、「構成」は 4 点満点、「言語使用」は 5 点満点とした。

(4) の Fluency の視点から選択した総語数と文数は本研究の CEFR-J A2 レベルの学習者によって記述された自由英作文においても説明変数候補としたものである。

(5) の文の視点から選択した 1 文の平均語数は本研究の CEFR-J A2 レベルの学習者によって記述された自由英作文においても説明変数候補としたものである。Error-free T-unit 平均語数を採用しなかったのは、(3)の「言語使用」をエラーに数に基づいて、算出したため、エラーに関する変数を重複させないためである。

(6) の語彙の視点からは選択した lexical variation (語彙の多様性) を表す GI、D、および MTL、lexical sophistication (語彙の広範さ) を表す P_Lex および S、平均語長は本研究の CEFR-J A2 レベルの学習者によって記述された自由英作文においても説明変数候補としたものである。

(7) のリーダビリティの視点から選択した FKG、ARI、CLI は本研究の CEFR-J A2 レベルの学習者によって記述された自由英作文においても説明変数候補としたものである。

本研究の CEFR-J A2 レベルの学習者によって記述された自由英作文の説明変数候補として選択した(3) 談話の Halliday and Hasan (1976)における結束性に基づく 1 英文テキストにおける結束数の合計、結束数の合計に占める指示 (R) および接続 (C) の割合は説明変数候補から除外した。結束性は初級学習者の自由英作文においては重要であるとされる

(橋内, 1995) が、CEFR-J B2 レベルおよび C1 レベルの上級学習者を対象とするため、結束性を説明変数候補とすることは控えた。当初は、談話の説明変数候補として、結束性に変わる首尾一貫性 (coherence)^{注18)} という特定要因の評価指標の作成を考えていた。しかし、この首尾一貫性は、本研究では「構成」および「内容」で作成した特定要因の評価指標とも重複すると考え、談話の視点からの指標という選択は差し控えることにした。

このようにして選択した特定要因の評価指標を説明変数候補として、自由英作文の全体的評価を目的変数とする重回帰分析を行い、CEFR-J B2 および C1 レベルの自由英作文の言語的特徴を分析する。この 7 つの視点からの説明変数候補であれば、重回帰分析の結果、決定係数 R^2 が最低限 0.50 を超えることは可能であると考えられる。

その重回帰分析の前段階として、CEFR-J A2 レベルの自由英作文の場合と同様に、CEFR-J B2 レベルおよび C1 レベルの自由英作文をリーダビリティと語彙の豊かさ (多様性と広範さ) の視点からもより詳細な分析を試みることにする。

前者のリーダビリティの分析は、CEFR-J B2 レベルおよび C1 レベルの自由英作文の言語的特徴を表す説明変数候補の特定要因の評価指標として、FKG、ARI、CLI の 3 つの指標でどれが適しているかの判断にも役立つと考えられる。また、新たな知見が得られれば、本研究の主目的である CEFR-J B2 レベルおよび C1 レベルの自由英作文のルーブリックのデザインの基礎データになりうるとも考えられる。

後者の語彙の豊かさ (多様性と広範さ) の視点から分析は、CEFR-J B2 レベルおよび C1 レベルの自由英作文の自由英作文の言語的特徴を表す説明変数候補の特定要因の評価指標として、語彙の多様性を表す指標である GI、D、MTLD ではどれが、また、語彙の広範さを表す指標である P_Lex と S ではどちらが、適しているのかの判断にも役立つと考えられる。また、新たな知見が得られれば、本研究の主目的である CEFR-J B2 レベルおよび C1 レベルの自由英作文のルーブリックのデザインの基礎データになりうるとも考えられる。

2.6 まとめ

自由英作文には以下の 3 つの評価手法がある。

- (1) 評価者の印象度にもとづく全体的評価
- (2) 自由英作文のテキストを文法、語彙、構成などの項目にわけ、それぞれの項目を評価し、その結果を合算する分析的評価
- (3) 例えば、自由英作文の書かれている文字の語数 (総語数) という特定の要因を基に評価する特定要因の評価

^{注18)} 米山 (2003: 41) によれば、「テキストの内容がバラバラでなく、まとまりがあって首尾一貫している状態であり、意味上のまとまりがあり、内容が全体としてつじつまが合う状態」のことを言う。それ故、「明示的な言語形式で表されず、読み手 (もしくは聞き手) が、常識的な現実認識を作用させ」(寺澤, 2002: 105)、テキストを理解する必要がある。

従来のルーブリック (As Is) は、全体的評価と分析的評価によって、作成されている。まず、全体的評価を使用したルーブリックには馬場(2008)と‘Independent WRITING Rubrics’を提示した。前者の馬場 (2008) は本研究で初級英語学習者とした CEFR-J A2 レベルよりも英語力が低いと推定される学習者を対象としている。後者は TOEFL の Independent 型 WRITING の評価で使用されるもので、上級から初級の学習者までを対象とする汎用性の高いルーブリックである。これら 2 つの全体的評価を使用したルーブリックについて長所と短所を分析した。

分析的評価を使用したルーブリックとして、ESL Composition Profile と Nakanishi(2006)による EFL Composition Profile、Pilot Rubric for Paragraph / Essay Writing を提示した。ESL Composition Profile は Jacobs et al(1981)によって作成され、第 2 言語として英語を学ぶ学習者と対象としている。EFL Composition Profile は第 2 言語でなく、外国語として英語を学ぶ日本人の大学生を対象に ESL Composition Profile を改良したものである。Pilot Rubric for Paragraph / Essay Writing はアメリカの 2 つのパブリックスクールの第 1 言語向けのルーブリックを結合した後に、日本人の大学生の自由英作文の評価基準となるように修正を加えたものである。これら 3 つの分析的評価を使用したルーブリックの長所と短所を分析した。

本研究のルーブリック (To Be) では、特定要因の評価手法に基づく分析的評価を採用するため、自由英作文の言語的特徴を重回帰分析によって抽出した説明変数を特定要因の評価指標とし、それをルーブリックの評価の観点とし、分析的評価の各項目とするという基本的な考え方を提示した。さらに、評価基準 (到達の度合い) に関しても実証された数値に基づいて、その水準を定めることとした。これらの手法は全体的評価や分析的評価を使用した既存のルーブリックには使用されていないものである。

次に、この自由英作文の言語的特徴に関する 4 つの先行研究を提示した。

- (1) Ferris(1994) : アメリカの大学に留学している CEFR-J B2 レベル以上の英語力と考えられる留学生によって記述された自由英作文を重回帰分析によって調査し、その言語的特徴は総語数、同義語・反意語、語長、受動態、3 人称・非人称代名詞となった。説明率の $R^2=0.50$ である。
- (2) 石田 (2013) : 英語母語話者と非母語話者 (英語力 CEFR-J B1 以上) の自由英作文を判別分析を用いて調査した結果、総語数、異なり語数、GI、2 語表現の総数が両者を判別する言語的特徴となった。
- (3) 杉浦 (2008) : CEFR-J B1 以上の英語力を有する日本人大学生および大学院生によって記述された自由英作文を重回帰分析によって調査した結果、総語数、文あたりの従属節数、文あたりの接続語句数、平均文長が言語的特徴となった。自由度修正済み決定係数は 0.44 であった。
- (4) 水本 (2008) : CEFR-J A2 レベルの下位の英語力を有する日本人大学生によって記述された自由英作文を重回帰分析によって調査した結果、総語数、FKG、平均語長が自

由英作文の言語的特徴となった。調整済み $R^2=0.60$ である。

これらの先行研究を考慮し、本研究においても、自由英作文の言語的特徴を調査するために、全体的評価を目的変数とする重回帰分析を使用し、決定係数 R^2 が最低限 0.50 を超えることを目処とし、特定要因の評価指標候補を選択することとする。

CEFR-J A2 レベルの学習者によって記述された自由英作文の特定要因の評価指標候補として、Fluency (総語数、総文数)、文 (1文の平均語数、Error-free T-unit 平均語数)、談話 (結束数の合計、結束数の合計に占める指示 (R) および接続 (C) の割合)、語彙 (GI、D, MTL D、P_Lex、S、平均語長)、リーダビリティ (FKG、ARI、CLI) の 5 つの視点から 15 変数を説明変数候補として選択することとした。

CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文特定要因の評価指標候補として、構成 (本研究で独自に作成した「構成」という特定要因の評価指標)、内容 (本研究で独自に作成した「内容」という特定要因の評価指標)、言語使用 (本研究で独自に作成した「言語使用」という特定要因の評価指標)、Fluency (総語数、文数)、文 (1文の平均語数)、語彙 (GI、D, MTL D、P_Lex、S、平均語長)、リーダビリティ (FKG、ARI、CLI) の 7 つの視点から、15 変数を説明変数候補として選択することとした。

3. CEFR-J A2 レベルの自由英作文の評価の観点の提案

3.1 CEFR-J A2 レベルの自由英作文のリーダビリティによる分析

3.1.1 はじめに

3章1節では、CEFR-J A2 レベルの学習者によって記述された自由英作文をリーダビリティの視点から分析する。リーダビリティは、小池（2003：63）による「リーダビリティとは、テキストの読みやすさ、理解しやすさ」という定義を採用する。このリーダビリティを表す様々な指標が考案され、ここでは、それらの代表的な指標である Flesch-Kincaid Grade（以下、FKG とする）、Automated Readability Index（以下、ARI とする）、Coleman-Liau Index（以下、CLI とする）を分析に使用する。これらの指標では、テキストを読むために最低限必要とされる教育レベルをアメリカの学年で表示される（染谷，2009）。例えば、あるテキストを FKG で分析した結果、9.0 という数値になれば、そのテキストを読みには最低限、アメリカの9年生の教育レベルに達していなければならないことになる。

このテキストの読みやすさを測定するリーダビリティを染谷（2009）では、自由英作文のようなライティングのテキストの分析に活用することを推奨している。しかしながら、日本ではリーダビリティをライティングに活用した研究があまりなされていないのが現状である。そのため、CEFR-J A2 レベルの初級学習者によって記述された自由英作文のリーダビリティは、どの程度になるのかも判明していない状況である。

そこで、本節では CEFR-J A2 レベルの自由英作文をリーダビリティの視点から分析し、本研究の主目的であるルーブリックの評価の観点としての自由英作文の言語的特徴としてリーダビリティの指標が有用であるかどうか、もし、有用であるならば、どの指標を用いるのが適切であるのかを調査する。

3.1.2 研究の背景

本節ではリーダビリティの定義として、小池（2003：63）を採用し、高梨・卯城（2000：30）は、リーダビリティについて「(1)手書き文字または活字の大小、種類、組み方等による見やすさ、(2)文章のおもしろさや楽しさによる読みやすさ、(3)語彙の難易、文長等の文体による読みやすさを挙げ、一般には(3)の意味で使用される」としており、ここでも、「(3)語彙の難易、文長等の文体による読みやすさ」に準拠してリーダビリティを扱うこととする。

このようなリーダビリティに関しては様々な公式が考案されてきており、小池（2003：63）では「公式の要素は主に難易度（語彙リスト、シラブル数）とセンテンスの長さ（具体的には語数、文法的複雑さの指標としてもとらえる）、学年、または年齢との調整をとるための定数からなる。リーダビリティ・スコアは学年あるいは年齢で算出される」としている。指標としてスタンダードで定評のあるのが FKG であり、対象とするテキストを読むために最低限必要とされる教育レベルをアメリカの Grade Level で示している。また、FKG はその算出にあたり、音節にも依拠しており、一方、音節に全

く依拠しない指標として、ARI や CLI がある（染谷，2009）。

これらのリーダビリティの公式は通常リーディングに関する研究に使われることが多い（例えば、水本，2013；Hamaoka, 1999；石岡・橋本・大津，2009 など多数）。しかしながら、この指標を自由英作文のようなライティングに活用した研究は日本ではあまり行われていないようである。以下に、その数少ないリーダビリティをライティングに活用した先行研究を紹介する。

樋口（2010）は国立大学教育学部の学生を対象とした共通教養科目の中のライティングの授業（15回）において、英文のテキストを分析する指標においてリーダビリティ値（Kincaid 値^{注19}）と Fog Index 値^{注20}）を使用している。対象とする人数は12名である。この内、リーダビリティ値が確実に向上したのが3名としている。学生BのKincaid 値は「3.8→4.4→4.9」、Fog Index 値は「7.0→7.7→7.9」となり、学生FのKincaid 値は「3.5→3.8→4.3」、Fog Index 値は「6.3→6.4→6.9」となり、学生HのKincaid 値は「3.0→2.3→3.1」、Fog Index 値は「5.6→4.8→6.0」となり、一旦は下降しても、当初より確実に向上している。国立大学の教養課程で英語を学ぶ学生の英語力は明確に示されていないが、おそらくCEFR-JのA2レベルの英語力ではないかと考えられる。付録資料として12名のKincaid 値とFog Index 値も提示されているが、人数が少なく、一般化するにはやや無理があると言えよう。

水本（2008）では、自由英作文における語彙の統計指標と全体的評価の関係を調査しており、その語彙の統計指標の1つとしてFKGが使用されている。このFKGは自由英作文の全体合的評価を目的変数とした説明変数として抽出されている。自由英作文の書き手はTOEICで300点前半程度の英語力の40名の大学生とされており、CEFR-JのA2レベルの範疇にはいるのではないかと考えられる^{注21}。FKGの実測値は提示されていないが、回帰式からの予測得点と実際の評価の残差が±1.5以上だった8名はFKGが7.2から3.9であることが記載されている。CEFR-J A2レベルの大学生の記述する自由英作文に関するFKGの値として参考になると言えよう。

これらの結果を踏まえて、CEFR-JのA2レベルの自由英作文をリーダビリティの視点から分析していくこととする。

3.1.3 研究の目的

本節では、CEFR-JのA2レベルの学習者が記述する自由英作文はリーダビリティの視点から分析し、そのテキストを読むために最低限必要とされる教育レベルはアメリカの学年で、どのくらいのレベルになるのかを明らかにし、そのリーダビリティの指標が本研究の

注19) Flesch-Kincaid Grade である。

注20) Robert Gunning が1952年に作成したもので、FKGよりやや高めの数値となる（染谷，2009）。

注21) TOEIC形式の模擬試験によって、推定された英語力である。

主目的であるルーブリックの評価の観点、言い換えれば、自由英作文の言語的特徴の説明変数候補として有用かどうか、もし、有用と考えられるならば、どの指標を用いるのが妥当なのかを調査することを目的とする。本研究で取り扱うリーダビリティの指標として、最も一般的な FKG、ARI、CLI を用いることとする。

以下に、この目的を達成するために 2 つの研究課題を設定する。

- (1) CEFR-J A2 レベルの学習者が記述した自由英作文のテキストでは、リーダビリティの指標である FKG、ARI、CLI の平均値はどうか。
- (2) リーダビリティの指標である FKG、ARI、CLI の中で、CEFR-J A2 レベルの学習者が記述した自由英作文を評価するのに、どの指標が最も妥当性が高いか。

3.1.4 方法

3.1.4.1 CEFR-J A2 レベルの調査参加者

CEFR-J A2 レベルの英語力を有する A 高等学校（中部圏に所在）の生徒 106 名（2 年生：48 名，3 年生：58 名）を調査参加者とした。全員が文科系の大学進学を希望している。表 3-1 はベネッセコーポレーション主催の英語コミュニケーション能力テストである GTEC for STUDENTS（以下、GTEC とする）による調査参加者の英語力に関する記述統計である。GTEC は RS、LS、WS より構成されている。各セクションの配点は RS が 320 点満点、LS が 320 点満点、WS が 160 点とされ、3 セクション合計の満点は 800 点である（この配点は 2005 年当時の GTEC であり、現在は配点に変更されている）。GTEC の得点は、「TOEIC の得点 = GTEC の得点 × 1.04 - 122」という変換式で、TOEIC の得点に換算可能である。

この 106 名は平林（2005）の 180 名から CEFR-J A2 レベルの生徒のみを抽出した。GTEC の得点を TOEIC に換算すると、平均点が 373 点となり、最小値は 242 点、最大値は 548 点となった。CEFR-J A2 レベルは TOEIC 225 点～550 点（河原，2013）とされるので、この調査参加者の英語力は CEFR-J A2 レベル全般をほぼ満たすと判断した。

表 3-1 調査参加者の英語力の記述統計（ $N = 106$ ）

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
R S	320	197.37	24.99	129	269	140	0.22	0.18
L S	320	179.09	32.60	87	320	233	1.11	4.33
WS	160	99.72	19.00	50	130	80	0.26	-0.87
総合点	800	476.18	59.05	350	644	294	0.42	0.67
TOEIC								
換算点	990	373.21	61.44	242	548	306	0.42	0.67

RS:リーディングセクション LS:リスニングセクション WS:ライティングセクシ

3.1.4.2 GTEC ライティング

自由英作文は GTEC で記述された英文テキストであり、専門評価者の全体的評価による印象度で、得点がつけられている。表 3-1 の WS が、その全体的評価の得点である。課題は「あなたはアメリカの現地校に短期留学をしていて、次のテーマで英文エッセイを提出することになりました。課題：現在、あなたの身の周りで起こっている事から日本の社会が抱えている問題点と思われる事例を 1 つ取り上げて、そのことについてあなた自身の考えを述べなさい」というものである。この課題に対して、調査参加者は 20 分で辞書や参考文献を利用することなく自由英作文を書くことになる。

3.1.4.3 リーダビリティの指標

リーダビリティの指標として使用する FKG、ARI、CLI について、染谷 (2009) をもとに、概略を以下に示す。リーダビリティの指標として様々な指標が提示されているが、これらは一般に対象テキストの語数、文数、音節数をもとに算出されるものである。FKG は米国の学年レベルに対応させたものであり、以下の公式によって算出される。

$$FKG=(0.39 \times \alpha) + (11.8 \times \beta) - 15.59$$

α = average number of words per sentence β = average number of syllables per word

この FKG は “authentic” な英文の場合、そのスコア 5.5 から 16.3 の範囲内におさまるとされる。

その算出の際に音節に依拠しないのが ARI と CLI である。以下の公式で算出される。

$$ARI=(4.71 \times CPW) + (0.5 \times WPS) - 21.43$$

$$CLI=(5.89 \times CPW) + (0.3 \times (100 \div WPS)) - 15.80$$

CPW=average number of characters per word

WPS=average number of words per sentence

ARI は米国空軍内の文書（特に技術文書やマニュアル）の作成・評価を支援する目的で提案され、コンピュータの自動処理を念頭に入れて考案されたものである。CLI も同様にコンピュータの自動処理を前提にしている。これらの結果も米国の学年レベルで算出される。ARI は FKG や CLI よりも高めに算出されことも Taylor (2008) が示している。

このような説明に加えて、染谷 (2009) では FKG、ARI、CLI に関して、英文サンプルをもとにした相関関係も分析している。結果として、FKG、ARI、CLI の相互には 0.9 を超える強い正の相関関係があることを提示している。

3.1.4.4 手続き

(1) 自由英作文のテキストデータの収集

平林 (2005) で収集した 180 名の自由英作文のテキスト (MS-Word によりデジタルデータ化済み) から、CEFR-J の A2 レベルの英語力の対象となる本研究の調査協力者を抽出する。

(2)リーダビリティの指標である FKG、ARI、CLI のスコアの算出

①FKG のスコアは MS-Word のスペルチェックと文章校正機能により算出する。

②ARI と CLI のスコアは前青山学院大学染谷泰正教授（現関西大学教授）の作成された Word Level Checker

（英文語彙難易度解析プログラム：<http://someya-net.com/wlc/index J.html>）

により算出する。

(3)FKG、ARI、CLI に関する記述統計量の算出

(4)GTEC の WS（自由英作文の全体的評価の得点）、FKG、ARI、CLI の相関関係の調査

3.1.5 結果と考察

表 3-2 はリーダビリティの指標に関する記述統計である。この結果から、平均値は CLI が最も高い 6.00、次に FKG の 4.22、最も低いのが ARI の 3.37 となった。この数値はアメリカの学年を示しているので、アメリカの小学校 3 年生レベルから 6 年生程度のリーダビリティということになる。

FKG は“authentic”な英文の場合、5.5 から 16.3 の範囲内に収まるとされるが、CEFR-J A2 レベルの自由英作文のテキストの場合、そのスコアは 4.22 であり、“authentic”な英文の範疇に入らないことが判明した。非母語話者で外国語として英語を学ぶ初級英語学習者の自由英作文の特性を示しているものと考えられる。

Taylor (2008) では、ARI は FKG や CLI よりも高めに算出されると指摘されているが、CEFR-J A2 レベルの調査参加者の自由英作文では、ARI が最も低い数値となるという結果を得た。ARI というのはもともと、技術文書の作成・評価を念頭に考案されているので、CEFR-J A2 レベルの自由英作文では、そのテキストは技術文書の水準に及ばないと考えるのが妥当で、かえって低いスコアとなったのではないかと考えられる。このことも非母語話者で外国語として英語を学ぶ初級英語学習者の自由英作文の特性と言えよう。

表 3-2 リーダビリティの指標に関する記述統計 (N = 106)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
FKG	4.22	1.36	1.40	7.80	6.40	0.32	0.19
ARI	3.37	1.77	-0.30	8.50	8.80	0.44	-0.14
CLI	6.00	1.91	1.90	11.10	9.20	0.42	-0.22

*ARI の最小値が負の値になっているのは、その算出式の中の CPW=average number of characters per word の値が低い為に、生じたと考えられる。ARI でこの数値となった自由英作文の CLI も最低値の 1.90 となっている。但し、算出の際に、音節に依拠する FKG は 2.10 と最小値とはなっていない。

表 3-3 は自由英作文の全体的評価と FKG、ARI、CLI の相関行列表である。自由英作文の全

体的評価とFKG、ARI、CLIのピアソンの相関係数は0.24、0.24、0.21であり、弱い正の相関関係があるという結果となった。弱い正の相関関係があるため、これらのリーダビリティの指標が、自由英作文の言語的特徴となる可能性はあると言えよう。

また、FKG、ARI、CLIの3指標間のピアソンの相関係数は0.87を上回り、ARIとCLI間にはピアソンの相関係数が0.96となり、強い正の相関関係があることが判明した。これは、染谷（2009）における英文サンプルにおけるFKG、ARI、CLI間のピアソンの相関係数0.9を超える強い正の相関関係があるという結果とほぼ同様の水準の相関係数となった。

表 3-3 自由英作文の全体的評価と FKG、ARI、CLI の相関行列表

	(1)	(2)	(3)	(4)
(1) WS (自由英作文の全体的評価)	—	.24*	.24*	.21*
(2) FKG		—	.87**	.87**
(3) ARI			—	.96**
(4) CLI				—

*: $p < .05$ **: $p < .01$

このような結果から、以下のように考察した。

研究課題(1)の「CEFR-J A2 レベルの学習者が記述した自由英作文のテキストはリーダビリティの指標である FKG、ARI、CLI の平均値はどうなるのか」の解答は、FKG、ARI、CLI がそれぞれ 4.22、3.37、6.00 となり、CEFR-J A2 レベルの自由英作文を読むためには、アメリカの小学校 3 年生レベルから 6 年生程度の最低限の教育レベルが必要なことが判明した。言い換えれば、非母語話者として外国語を学ぶ日本人の高校生の自由英作文は母語であってアメリカ人の小学校 2 年生以下では読むことが難しいということである。

研究課題(2)の「リーダビリティの指標である FKG、ARI、CLI の中で、CEFR-J A2 レベルの学習者が記述した自由英作文を評価するのに、どの指標が最も妥当性が高いか」の解答は以下のように考察した。表 4-3 に示したように、FKG、ARI、CLI のいずれの指標も、自由英作文の全体的評価とは弱い正の相関関係がある。このことはリーダビリティが自由英作文の言語的特徴を表す指標として有用になる可能性を示唆している。また、3 指標間には相関係数 0.8 以上^{注22)} の強い正の相関関係があるため、どの指標を選択することも可能であるとも考えられる。しかしながら、CEFR-J A2 レベルの自由英作文では、FKG が 4.22 となり、“authentic”な英文の 5.5 から 16.3 の範囲 (FKG の本来の測定対象領域と考えられる) を逸脱していることから、FKG の使用は控えることとした。また、ARI が本来であれば、FKG や CLI よりも高めの数値となるべきにもかかわらず、3 変数の中で最も低い数値

^{注22)} 平井（2012）によれば、重回帰分析の前提として、多重共線性の検討の必要があり、そのためには説明変数候補の変数間の相関係数 0.8 以上を避けるべきであるとしている。

の 3.37 になってしまっているため、CEFR-J A2 レベルの自由英作文の評価する指標として、ARI の使用も控えることとした。CLI はこの 3 指標の中では、最も数値の高い 6.0 という結果となっており、この数値は CEFR-J A2 レベルの高校生の学年に最も近い数値であることも考慮した。従って、研究課題(2)の解答として、CLI を CEFR-J の A2 レベルの学習者が記述した自由英作文を評価するリーダビリティの指標として、最も妥当性が高いと判断した。

3.1.6 結論

3 章 1 節では、CEFR-J A2 レベルの自由英作文をテキストの読みやすさおよび理解しやすさと定義したリーダビリティの指標である FKG、ARI、CLI により分析した。従来、リーダビリティはリーディングの研究に使用されることが一般的で、自由英作文のようなライティングにリーダビリティが活用された研究はあまり行われてこなかった。そのため、CEFR-J A2 レベルのような初級英語学習者によって記述された自由英作文とリーダビリティの関係は判明していなかったが、以下の事項が明らかとなった。

- (1) CEFR-J A2 レベルの自由英作文を読むためには、アメリカの小学校 3 年生レベルから 6 年生程度の最低限の教育レベルが必要なこと。
- (2) リーダビリティが CEFR-J A2 レベルの自由英作文の言語的特徴を表す指標として有用になる可能性があること。
- (3) CEFR-J A2 レベルの自由英作文を評価するリーダビリティの指標は CLI が妥当であること。

この結論に基づき、2 章 5 節 2 項の「CEFR-J A2 レベルの自由英作文の評価の観点となる特定要因の評価指標候補」のリーダビリティの視点の説明変数候補から、FKG と ARI を除外することとした。

3.2 CEFR-J A2 レベルの自由英作文の語彙の豊かさの観点による分析

3.2.1 はじめに

近年、日本においても自由英作文の語彙の研究（小島，2013；石川，2005；水本，2008 など）がなされてきているが、CEFR-J A2 レベルの英語力に限定した学習者、とりわけ、その英語力の最も一般的と考えられる高校生によって書かれた自由英作文の語彙の豊かさに関する研究はほとんど行われていない。

海外における研究（Linnarud, 1986 など）では、自由英作文における語彙の豊かさはそのテキストの評価と関係があることを示唆している。このことは、語彙の豊かさを表す指標は自由英作文の言語的特徴となり、ルーブリックの評価の観点となりうる可能性があることを示している。

また、CEFR-J A2 レベルの学習者によって記述された自由英作文の語彙の豊かさから新たな知見が得られれば、自由英作文の評価基準であるルーブリックの基礎データとなりうると思われる。

3.2.2 研究の背景

コンピュータ技術の発達に伴い学習者コーパスの構築も進み、ライティングやスピーキングにおいて学習者が産出する発表語彙に関する研究が日本でも盛んに行われるようになってきている。本研究ではライティングの自由英作文の中で使用される発表語彙をその対象とする。

この発表語彙は学習者である書き手の有する語彙の豊かさ（lexical richness）という視点から研究が行われことがある。本研究では、この語彙の豊かさという視点から分析を進める。語彙の豊かさを分析する観点の主なものを以下に紹介する。Laufer（1994）や Laufer & Nation（1995）では語彙の豊かさを測定する一般的な方法として、①lexical originality（語彙の独創性）、②lexical density（語彙の密度）、③lexical sophistication（語彙の広範さ）、④lexical variation（語彙の多様性）を挙げ、それらの欠点を指摘するとともに、⑤Lexical Frequency Profile（LFP）を提唱し、その妥当性も検証している。さらに Engber（1995）などにもみられる⑥number of errors としての語彙の間違いの観点もある。

⑤の LFP に関して、海外の研究ではよく使用されている（Muncie, 2002; Lemmouh, 2008; Nadarajan, 2011 など）が、日本の一般的な高校生や大学生を対象とした研究ではあまり使用されないようである。この理由として、葉田野・馬場（2006）は平均的な日本人大学生の英作文では General Service List (West, 1953) の高頻度 1000 語の使用が多く、LFP は語彙の頻度別割合によって算出されるため（相沢, 2001）、言い換えれば、General Service List の低頻度語彙が作文に含まれていないと LFP による分析が困難であることを指摘している。また、望月他（2003）では、LFP は作文の長さが 400 語以上の比較的長めの作文でなければ処理できないという問題点を挙げている。この 2 つの指摘は CEFR-J A2 レベルの

英語学習者が書く自由英作文にも該当すると考えられるため、本研究では LFP の観点からの分析は困難となる。

語彙の豊かさに関する研究でライティングの評価との関係を分析したものに Linnarud (1986)、Engber (1995)、Daller & Phelan (2007) などがある。Linnarud (1986) では lexical individuality (lexical originality の変形) という指標が最も評価と相関関係 ($r = .47$) があり、Engber (1995) では lexical variation without error という指標が最も評価との相関関係 ($r = .57$) があり、次いで lexical variation with error が評価との相関関係が高くなっている ($r = .45$)。また、Daller & Phelan (2007) では、語彙の多様性を表す指標 GI が評価と最も高い相関関係 ($r = .57$) があるとしており、語彙の広範さを表す P_Lex と相関関係 ($r = .49$) があるとしている。いずれも中程度の正の相関関係を示している。これらの結果は語彙の豊かさがその評価を決定する際の重要な要因になりうることを示唆するものである。

Engber (1995) や Daller & Phelan (2007) のように多様性^{注23)} と広範さ^{注24)} という語彙の豊かさの視点から自由英作文の語彙を分析した研究に小島 (2013) がある。日本人英語学習者である大学生および大学院生 (CEFR-J B1 レベルおよび B2 レベルの英語力) と母語話者によって記述された自由英作文を分析している。多様性の指標として、GI、D、MTLD、広範さの指標として、P_Lex、S を使用している。結果として、語彙の多様性も広範さも英語母語話者とは明確に異なること、語彙の多様性と広範さは中程度の相関関係があるが、多様性に関しては習熟度の区別による差はなく、広範さに関しては、習熟度により差があることを明らかにしている。但し、この小島 (2013) では語彙の豊かさと自由英作文の全体的評価の関係は分析されていない。

本節では、語彙の豊かさに関する指標を用いて CEFR-J A2 レベルの自由英作文の語彙の豊かさについて調査を試みる。さらに、これらの指標に加えて、『CEFR-J Wordlist Version 1.0』(東京外国語大学投野由紀夫研究室) という語彙リストを本研究における語彙の豊かさの調査に活用する。投野 (2013 : 132) によると、「A1 (1165 項目)、A2 (1416 項目)、B1 (2453 項目)、B2 (2789 項目) の合計 7823 項目となっており、A2 レベルまでで 2000 語習得というのは、ネイティブ・スピーカーの言語使用をみても会話で 90%、書き言葉で 80% をカバーする語彙力なので、A2 レベルで約 1000 語程度が発表語彙として、残り 1000 語が受容語彙として使えれば十分である」と指摘している。同氏によれば、日本の学校教育での相当レベルとして A1 は小学校～中学校 2 年程度、A2 は中学 3 年～高校 1/2 年程度、B1 は高校 3 年～大学受験レベル、B2 は大学受験～大学教養レベルとされる。CEFR-J A2 レベ

注23) テキストにおける異語数と総語数に基づき、どの程度多様な語彙が使用されているかを評価することを表し、指標として、TTR (type-token ratio)、GI、D、MTLD などがある (小島, 2013)。

注24) 語彙の一般的な頻度情報を利用して、頻度の観点から自由英作文の語彙使用がどの程度広範なものかを評価することを表し、指標として、P_Lex、S などがある (小島, 2013)。

ルの自由英作文では A1 レベルの高頻度語彙のどの程度、また、B2 レベルのような低頻度語彙はどの程度使用されるのかを分析する。

3.2.3 研究の目的

先行研究より発表語彙の豊かさが自由英作文の評価を決定する要因になりうることを示唆している。3章2節の分析では、CEFR-J A2 レベルの学習者が記述する自由英作文の発表語彙の豊かさを多様性と広範さの観点から分析し、その豊かさが自由英作文の質に影響するかどうか、また、多様性と広範さにおいてどのような使われ方をしているかを明らかにすることを目的とする。この目的のために以下の研究課題を設定する。

- (1) CEFR-J A2 レベルの自由英作文の全体的評価と発表語彙の豊かさ（多様性と広範さ）の関係およびその多様性と広範さの関係はどのようになっているのか。
- (2) 英語の習熟度（3群に分ける）が異なると、多様性と広範さの各指標のスコアには差異があるのか。
- (3) CEFR-J A2 レベルの学習者が記述した自由英作文では、CEFR-J の語彙リスト『CEFR-J Wordlist Version 1.0』のどのレベルの語彙を使用しているのか。
- (4) 『CEFR-J Wordlist Version 1.0』では英語の習熟度の違いにより、語彙レベルの使われ方、特に低頻度語彙には差があるのか。

3.2.4 研究方法

3.2.4.1 CEFR-J A2 レベルの調査参加者

調査参加者は3章1節と同一である（表 3-1 参照）。英語の習熟度の相違により自由英作文の語彙の使われ方を調査するために、CEFR-J A2 レベルの英語力を上位者・中位者・下位者の3段階に分けることとした。その分類は以下のようにして行った。Educational Testing Service (ETS) のホームページに掲載されている TOEIC の得点（第 86 回公開テスト）と英検取得者の関係、すなわち、英検 2 級・準 2 級・3 級の各級の取得者の最も多いのが、それぞれ 445 点～495 点、345 点～395 点、245 点～295 点というデータを参考にし、上位者 35 人（548 点～401 点）、中位者 33 人（400 点～351 点）、下位者 38 人（350 点～242 点）に分類した。この上位者・中位者・下位者の英語力の記述統計を表 3-4 に示した。

表 3-4 上位者・中位者・下位者の英語力（TOEIC 換算）の記述統計

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
上位者 (N=35)	990	441.20	39.61	401	548	147	1.16	0.67
中位者 (N=33)	990	373.23	16.16	350	397	47	0.11	-1.63
下位者 (N=38)	990	310.61	25.27	242	348	106	-0.91	0.50

3.2.4.2 GTEC ライティング

GTEC ライティングは 3 章 1 節と同一である。

3.2.4.3 語彙の豊さの指標について

3 章 2 節では、語彙の豊かさを測定する指標に小島（2013）で使用された多様性の指標の GI、D、MTLD、広範さの指標の P_Lex、S の 5 つの指標について以下に説明する。

(1)GI は Guiraund Index の略語である。語彙の多様性とは、4 章の脚注 5 に記述したように「テキストにおける異語数と総語数に基づき、どの程度多様な語彙が使用されているかを評価すること」を表し、様々な研究で使用されるのが TTR (type-token ratio) であり、テキスト中の総語数 (token) に対する異語数 (type) の割合を意味し、数値が高ければ、そのテキストの内容が濃く、また難度が高いと解釈される (米山, 2003)。しかし、この TTR は総語数の増加に伴い、その数値が下がるという問題がある。これはテキスト内で繰り返し使用される語が増え、総語数が増加すると、異語数の増加率が徐々に減少するためである。この問題点を解決するために、様々な多様性の指標が提案されている。その一つが GI である。GI では、TTR の分母を総語数の平方根とする修正式によって算出される指標である (小島, 2013)。しかしながら、この修正式を用いても尚、総語数の影響を受けるという欠点があるため、D や MTLD が開発されている。

(2)D は TTR の問題を解決するために、Malvern and Richards (1997) によって提案されたものである。この D は TTR 曲線がいずれも下降曲線を描くが、語彙が多様であるほど下降がゆるやかになり、TTR 曲線が上部に位置するという性質をもとにしており、そのスコアは TTR 曲線がどの程度、上部に位置するかを示すものであり、スコアが高いほど、語彙が多様である (小島, 2013)。D は free software である D_Tools によって算出できる。

(3)MTLD は the Measurement of Textual Lexical Diversity の略語である。この MTLD は McCarthy & Jarvis (2010) により開発され、D よりもテキストの長さ (総語数) の影響を受けにくいとされる。テキスト内の語彙の多様性を示す指標であり、異語数の増加が停滞するポイント (TTR=0.72) までに要するテキストの平均語数から算出され、このスコアが高いほど、語彙が多様であると解釈される (McCarthy & Jarvis, 2010; 小島, 2013)。

この MTLD は Web 上の (<http://gnutiez.de/wp/apps/mtld/>) プログラムによって値を算出できる。

(4) P_Lex は Meara & Bell (2001) によって開発された語彙の広範さを表す指標である。この P_Lex では、低頻度語の出現確率を利用し、一般的なスコアは 0~4.5 とされ、スコアが高いほど低頻度語彙の使用確率が高いことを示す (小島, 2013)。P_Lex は free software である分析ツールによって算出できる。

(5) S は小島 (2011) によって開発された語彙の広範さを表す指標である。テキストから 50 語ずつ取った各サンプルにおける高頻度語彙の累積カバー率をデータとし、データに最も近似するモデルを求め、累積カバー率が 100% に達する単語の頻度順位を推定する。例えば、S が 2016 の場合、そのテキストの発表語彙レベルは 2016 語と解釈される (小島, 2013)。S への Web 上の (http://www.kojima-vlab.org/S_J.html) プログラムによって値を算出できる。

3.2.4.4 語彙の豊さの指標の記述統計

3 章 2 節の研究では、語彙の豊さを測定する指標に関して、小島 (2013) で使用された GI、D、MTLD、P_Lex、S の 5 つの指標を用いることを当初考えていた。しかしながら、GI は総語数の影響を受けることを考慮して、ここでは多様性の指標とすることは控えた。また、同じく多様性の指標である D に関しては、対象となる英文テキストが 50 語以上から構成されている必要があり、本研究の調査協力者 106 名の中にも、3 名が 50 語に至らない英文テキストを書いていた。CEFR-J A2 レベルの英語力の場合、50 語を一定時間 (20 分) に書くことができないこともありうると判断し、D という多様性を表す指標も分析の対象からはずした。このため、2 章 5 節 2 項の CEFR-J A2 レベルの自由英作文の評価の観点となる特定要因の評価指標候補から、語彙の多様性を表す指標である GI と D を除外することとする。以下の表 3-5 に各指標に関する記述統計を示す。

表 3-5 語彙の豊かさの指標に関する記述統計

N=106	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
MTLD	71.55	24.81	29.60	166.06	136.46	1.01	1.19
P_Lex	0.17	0.18	0.02	0.72	0.70	1.29	1.22
S	2853.59	1327.97	765.00	7142.00	6377.00	0.75	0.73

3.2.4.5 手続き

(1) 自由英作文のテキストデータの収集

平林 (2005) で収集した 180 名の自由英作文のテキスト (MS-Word によりデジタルデータ化済み) から、CEFR-J A2 レベルの英語力の対象となるテキストを抽出した。

(2) 語彙の多様性の指標である MTLD の算出

MTLD は Web 上の (<http://gnutiez.de/wp/apps/mtld/>) のプログラムにより算出した。

(3) 語彙の広範さの指標である P_Lex、S の算出

P_Lex は分析ツール (<http://www.lognostics.co.uk/tools/>からダウンロード) を使用して算出した。

S は S の分析ツール (http://www.kojima-vlab.org/S_J.html) により算出した。

(4) 語彙の豊かさに関する各指標と自由英作文の全体的評価と相関関係の調査

(5) 習熟度の異なる 3 群において、各指標の平均値の有意差の検証

(6) CEFR-J A2 レベルの学習者が記述した自由英作文の発表語彙が、『CEFR-J Wordlist Version 1.0』のどのレベルの語彙を使用しているかの調査

3.2.5 結果と考察

3.2.5.1 自由英作文の全体的評価と語彙の豊かさに関する指標の関係

表 3-6 は自由英作文の全体的評価と語彙の豊かさに関する指標の相関行列表である。この結果より、自由英作文の全体的評価と語彙の多様性を表す指標との関係に関して、MTLD とは中程度の正の相関関係 ($r=.37$) があることが判明した。自由英作文の全体的評価と語彙の広範さを表す指標との関係に関して、P_Lex では弱い正の相関関係 ($r=.23$) があるが、S ではほとんど相関関係がないことが判明した。また、P_Lex と S はほとんど相関関係がないことも明らかとなった。語彙の多様性と語彙の広範さの関係に関しては、MTLD と P_Lex は弱い正の相関関係 ($r=.22$) があるが、MTLD と S はほとんど相関関係がないことが明らかとなった。

従って、研究課題(1)「CEFR-J A2 レベルの自由英作文の全体的評価と発表語彙の豊かさ(多様性と広範さ)の関係およびその多様性と広範さの関係はどのようになっているのか」の解答は以下ようになる。全体的評価と語彙の多様性に関しては、CEFR-J A2 レベルの自由英作文では、中程度の正の相関関係があるとする Engber (1995) と同様の結果となった。一方、全体的評価と語彙の広範さとは弱い正の相関関係に留まった。

多様性と広範さの関係に関しては、小島 (2013) や Daller & Phelan (2007) のように語彙の多様性と広範さと中程度の正の相関関係があるという結果ではなく、両者には弱い正の相関関係に留まり、両者の重なる部分は少ないことも明らかになった。これらの結果は、CEFR-J A2 レベルにおいてはメンタルレキシコン^{注25)}が十分に発達していないのが原因なのではないかと考えられる。

この分析で使用した広範さの指標である S は自由英作文の全体的評価とも広範さの指標である P_Lex とともに、さらに多様性の指標である MTLD とともにほとんど相関関係がないことが判明した。S のスコアがこのようになったのは、S の分析ツールで安定したスコアとなるには 200 語以上のテキストであることが求められるとあるため、本研究の 106 個の自由英作文のテク

注25) 心的語彙項目と日本語では訳され、頭脳に記憶として貯蔵されている語や語に関する情報 (米山, 2003)。

ストの総語数が最高値でも153語しかなく、そのスコアが不安定になったと思われる。このため、3章2節のCEFR-J A2レベルの自由英作文の評価の観点となる特定要因の評価指標候補から、語彙の広範さを表す指標であるSを除外することとした。

表 3-6 自由英作文の全体的評価と語彙の豊かさに関する指標の相関行列表

	Writing	MTLD	P_ Lex	S
Writing	--	.37**	.23*	-.03
MTLD		--	.22*	.07
P_ Lex			--	-.06
S				--

Writing: 自由英作文の全体的評価 * : $p < .05$ ** : $p < .01$

3.2.5.2 習熟度による語彙の豊かさの指標のスコア比較

表 3-7 は習熟度による語彙の豊かさの指標のスコアを比較したものである。指標ごとに一元配置の分散分析を行った。MTLD は習熟度の差により、有意差が認められた ($F(2, 103)=9.48, p < .01, \eta^2=.03$)。P_Lex は習熟度の差により、有意差は認められなかった ($F(2, 103)=1.61, p > .1$)。S は習熟度の差により、有意差は認められなかった ($F(2, 103)=1.73, p > .1$)。

MTLD に関して、チューキーの方法により多重比較を行った。結果として、MTLD では、上位者と中位者、上位者と下位者には有意差があった ($p < .01$) もの、中位者と下位者には有意差がなかった。

従って、研究課題(2)「英語の習熟度(3群に分ける)が異なると、多様性と広範さの各指標のスコアには差異があるのか」の解答は以下のようなようになる。語彙の多様性の観点では TOEIC 400 点の英語力の上下において語彙の語彙使用は区別された。一方、語彙の広範さに関しては習熟度による語彙使用は区別されなかった。これは、小島(2013)の習熟度の違いにより、多様性は区別されず、広範さにおいて区別される結果と相反することになった。小島(2013)のような CEFR-J B1 および B2 レベルの英語力とのメンタルレキシコンの発達の相違を反映しているのではないかと考えられる。つまり、CEFR-J B1 および B2 レベルではより低頻度の発表語彙を有しており、習熟度の違いにより、広範さが反映されることになるが、CEFR-J A2 レベルでは低頻度の発表語彙をあまり有していないため習熟度が異なっても、それを反映することはないのではないかと考えられる。

表 3-7 習熟度による語彙の豊かさの指標のスコアの相違

	上位者 N=35	中位者 N=33	下位者 N=38	多重比較
	M(SD)	M(SD)	M(SD)	
MTLD	85.20 (28.84)	67.11 (19.65)	62.84 (19.43)	**上位者 > 中位者、 **上位者 > 下位者
P_Lex	0.21 (0.21)	0.16 (0.18)	0.14 (0.15)	
S	2874.74 (1171.26)	2531.12 (1380.52)	3114.16 (1390.21)	

**： $p < .01$

尚、S の数値も参考までに掲載した。

3.2.5.3 『CEFR-J Wordlist Version 1.0』の語彙レベルからの分析

表 3-8 に習熟度の違いによる語彙レベル別使用率・中低頻度語彙および低頻度語彙使用率・レベルインデックスの各平均値を示した。この表では石川（2005）で使用されている手法に基づき、『CEFR-J Wordlist Version 1.0』での中低頻度語彙使用率（A2-B2 占有率）、低頻度語彙使用率（B1-2 占有率）、レベルインデックスを算出した。

研究課題(3)「CEFR-J A2 レベルの学習者が記述した自由英作文では、CEFR-J の語彙リスト『CEFR-J Wordlist Version 1.0』のどのレベルの語彙を使用しているのか」の解答を以下に示す。いずれの習熟度においても A1（小学校～中学校 2 年程度）が 90%程度を占めており、A2 レベル（中学校 3 年～高校 1/2 年程度）は 7%程度である。葉田野・馬場（2006）が平均的な日本人大学生の英作文では General Service List (West, 1953) の高頻度 1000 語の使用が多いことを指摘しているが、ワードリストは異なるものの、CEFR-J A2 レベルの学習者でも高頻度語彙の 1000 語程度までの語彙の使用が圧倒的である。B1 や B2 のような低頻度語彙は上位者でも 3%程度である。研究課題(1)での、語彙の広範さが多様性に比較して全体的評価に反映しないのは、低頻度語彙の発表語彙を有していない、言い換えれば、メンタルレキシコンが十分に発達していないと言えるのではないかと記述したが、それはこの語彙リストのレベル語彙使用率からも断定しても差し支えないようである。

英語の習熟度の違いにより中低頻度語彙および低頻度語彙使用率、レベルインデックスの各指標に差があるかどうか調べるために、指標ごとに一元配置の分散分析を行った。中低頻度語彙に関しては習熟度の差により、有意傾向が認められた ($F=(2, 103)=2.56, .05 < p < .01$)。低頻度語彙に関しては習熟度の差により、有意差は認められなかった ($F=(2, 103)=2.35, p > .01$)。レベルインデックスに関しては習熟度の差により、有意傾向が認められた ($F=(2, 103)=2.96, .05 < p < .01$)。

研究課題(4)「『CEFR-J Wordlist Version 1.0』では英語の習熟度の違いにより、語彙レベルの使われ方、特に低頻度語彙には差があるのか」の解答を以下に示す。英語の習熟度異なっても、低頻度語彙の使用の区別はされないことが判明した。中低頻度語彙や高頻度および中頻度の割合を含むレベルインデックスでは、習熟度の差により、語彙の使用に関して、有意傾向にあるに留まった。このことは研究課題(2)における語彙の広範さの指標で、低頻度語の出現確率を利用した P_Lex が習熟度の差により、有意差は認められなかったことと一致する。

表 3-8 習熟度による語彙レベル別使用率・中低頻度語彙および低頻度語彙使用率・レベルインデックスの各平均値

	A1	A2	B1	B2	A2+B1+B2	B1+B2	Lev Index
上位者	88.88	7.55	2.90	0.67	11.00	3.51	114.90
中位者	91.06	6.92	2.00	0.02	9.19	2.32	111.92
下位者	91.66	6.28	1.87	0.19	8.34	2.12	110.61

3.2.6 結論

3章2節では、CEFR-J A2レベルの英語力の学習者によって記述された自由英作文を語彙の豊かさを多様性と広範さという観点から分析した。従来、CEFR-J A2レベルの日本人学習者によって記述された自由英作文の語彙の分析をほとんど行われてこなかった。そのため、CEFR-J A2レベルのような初級英語学習者によって記述された自由英作文の語彙の豊かさは判明していなかったが、以下の事項が明らかとなった。

- (1) 語彙の多様性も語彙の広範さも自由英作文全体的評価とは相関関係がみられ、自由英作文の言語的特徴となることが示唆される。
- (2) 語彙の多様性の方が、語彙の広範さよりも自由英作文の全体的評価に関わりが強いことが判明した。
- (3) 語彙の多様性と語彙の広範さには弱い正の相関関係があり、両者の重なる部分は少ないことが明らかとなった。このことはCEFR-J A2レベルのような初級英語学習者のメンタルレキシコンの未発達の原因と考えられる。
- (4) CEFR-J A2レベル内での英語の習熟度の違いにより、語彙の多様性に関しては、TOEIC 400点を閾値として、語彙の使用は区別されることが明らかとなった。
- (5) CEFR-J A2レベル内での英語の習熟度が高くなると、語彙の多様性はより明確になるが、語彙の広範さは習熟度に影響されないことが明らかとなった。
- (6) CEFR-J A2レベルの学習者によって記述された自由英作文を『CEFR-J Wordlist Version 1.0』を基に語彙レベルを調査した結果、高頻度語彙に90%程度集中していることが明らかとなった。

これらの分析の過程で、語彙の多様性に関する指標である GI、D と語彙の広範さに関する指標である S は CEFR-J A2 レベルの自由英作文の語彙の調査には適切でないことも判明した。従って、2 章 5 節 2 項の「CEFR-J A2 レベル自由英作文の評価の観点となる特定要因の評価指標候補」の語彙の視点の説明変数候補から、GI、D、S を除外することとした。

3.3 重回帰分析により抽出した CEFR-J A2 レベルの自由英作文の評価の観点

3.3.1 はじめに

3章3節では、本研究の主目的たる CEFR-J A2 レベルの自由英作文のルーブリックにおける評価の観点と評価基準（到達度合い）の基礎データを決定する。評価の観点は、CEFR-J A2 レベルの自由英作文の言語的特徴として、重回帰分析によって抽出された説明変数である特定要因の評価指標を使用する。評価基準（到達度合い）は、説明変数となった特定要因の評価指標と自由英作文の全体的評価を基に統計的に検証した定量的数値により、3段階に区分することとする。

3.3.2 研究の目的

3章3節の目的は、CEFR-J A2 レベルの学習者によって記述された自由英作文の言語的特徴と自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値を明らかにすることである。この目的のために以下の研究課題を設定する。

(1) CEFR-J A2 レベルの学習者が記述する自由英作文の言語的特徴として有効な説明変数は何か。

(2) 自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値は何か。

(1)によって抽出された言語的特徴として有効な説明変数は CEFR-J A2 レベルの自由英作文のルーブリックにおける評価の観点となり、(2)の定量的な数値は評価基準（到達度合い）を決定する基礎データとなる。

本研究では、1つの課題の自由英作文で言語的特徴である説明変数を分析していることに疑問を抱かれることを考慮し、統計的手法を駆使することを徹底し、出来る限り一般化することに注意を払った。

3.3.3 方法

3.3.3.1 CEFR-J A2 レベルの調査参加者

調査参加者は3章1節と同一である（表3-1参照）。

3.3.3.2 GTEC ライティング

GTEC ライティングは3章1節と同一である。

3.3.3.3 CEFR-J A2 レベル自由英作文の評価の観点となる特定要因の評価指標候補

CEFR-J A2 レベルの自由英作文の言語的特徴に関する特定要因の評価指標の選択に関しては、先行研究を参考にしつつ、Fluency、文、談話、リーダビリティ、語彙の5つの視点から15変数とした（2章5節2項参照）。しかしながら、3章1節の結果から、CEFR-J A2 レベルの自由英作文を評価するリーダビリティの指標は CLI が妥当とし、FKG と ARI を除

くこととした。また、3章2節の結果から、語彙の多様性に関する指標であるGIとD、語彙の広範さに関する指標であるSはCEFR-J A2レベルの自由英作文の語彙の調査には適切でないことが判明したため、除外することとした。従って、CEFR-J A2レベル自由英作文の評価の観点となる特定要因の評価指標候補はFluency、文、談話、リーダビリティ、語彙の5つの視点から、以下の10変数となった。

- (1) Fluencyの視点からは総語数および文数を説明変数候補として選択した。
- (2) 文の視点からは1文の平均語数、Error-free T-unit 平均語数を説明変数候補として選択した。
- (3) 談話の視点からはHalliday and Hasan (1976)における結束性に基づく1英文テキストにおける結束数の合計、結束数の合計に占める指示(R)および接続(C)の割合を説明変数候補として選択した。
- (4) 語彙の視点からはlexical variation (語彙の多様性)を表すMTLD、lexical sophistication (語彙の広範さ)を表すP_Lex、平均語長を説明変数候補として選択した。
- (5) リーダビリティの視点からCLIを説明変数候補として選択した。

表3-10は言語的特徴に関する説明変数候補の記述統計を示す。

表3-10 CEFR-J A2レベルの自由英作文の言語的特徴に関する特定要因の評価指標の記述統計 (N = 106)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
総語数	98.70	29.67	43.00	174.00	131.00	0.52	-0.34
文数	10.42	3.39	4.00	22.00	18.00	0.87	0.86
1文平均語数	9.43	2.00	4.44	17.00	12.56	0.91	2.48
Error-free							
T-unit 平均語数	7.78	1.68	3.40	12.40	9.00	0.57	0.12
結束数の合計	8.87	3.71	1.00	19.00	18.00	0.33	-0.32
R+Cの割合	0.78	0.18	0.20	1.00	0.80	-0.72	0.39
MTLD	71.55	24.81	29.60	166.06	136.46	1.01	1.19
P_Lex	0.17	0.18	0.02	0.72	0.70	1.29	1.22
平均語長	4.38	0.40	3.51	7.15	3.64	3.17	19.97
数値変化後							
の平均語長	0.64	0.04	0.55	0.85	0.30	1.98	10.60
C L I	6.00	1.91	1.90	11.10	9.20	0.42	-0.22

表 3-10 の変数で、正規性の検定^{注26)}により平均語長^{注27)}に関しては正規性に関して問題があったため、数値変換を行った。

3.3.3.4 手続き

(1) 自由英作文のテキストデータの収集

平林 (2005) で収集した 180 名の自由英作文のテキスト (MS-Word によりデジタルデータ化してある) から、CEFR-J A2 レベルの英語力の対象となる本研究の調査協力者を抽出した。

(2) 言語的特徴を表す要因の選択とその算出

自由英作文の全体的評価 (GTEC の Writing Section の得点のため、以下 WS とする) に対する説明変数候補を Fluency (総語数、文数)、文 (1 文平均語数、Error-free T-unit 平均語数)、談話 (結束数の合計、結束数の合計に占める指示 (R) および接続 (C) の割合)、語彙 (MTLD、P_Lex、平均語長)、リーダビリティ (CLI) の 5 つの視点から 10 変数を選択し、数値を算出する。

(3) 重回帰分析

(2) の 10 変数を説明変数候補とし、自由英作文の WS を従属変数として、重回帰分析を行う。

(4) ① (3) で得られた各説明変数に関して、自由英作文の WS の上位者 (100 点～130 点:60 名)、下位者 (90 点～50 点:46 名) の 2 群に分け、その平均値に有意差があるかどうか調査する。

② (4) ①の平均値に有意差があれば、各説明変数に関して以下のように仮に定義する。

1. 上位者の平均値以上の数値をその説明変数に関する仮の上位者の定量的な数値とする。
2. 上位者と下位者の平均値の中間の数値をその説明変数に関する仮の中位者の定量的な数値とする。
3. 下位者の平均値以下の数値をその説明変数に関する仮の下位者の定量的な数値とする。

(5) (4) ②の仮の定量的な数値を以下のように検証する。

注26) 正規性については以下に示すような基準をもとに判断を下すことにした。(1) Kolmogorov-Smirnov の正規性検定 (2) 歪度/標準誤差 = ± 1.96 以下 尖度/標準誤差 = ± 1.96 以下 (3) 歪度・尖度が 1 を超えていない。(4) プロット化して分布が双峰分布になっていないかを確認した。

注27) 平均語長に関しては、プロット化した結果、L 字型分布はしており、双方分布や 矩形分布ではないため、正規分布の「きざし」はあるが(田中・山際, 1992)、尖度の値に問題があると判断し、数値変換を行った。数値変換には対数変換を行った。

- ①2 条件（自由英作文の WS の上位者と下位者）×3 値（各説明変数の上位者、中位者、下位者）の人数に関するクロス集計表を作成する。
- ②（5）①のクロス集計表を基に、 χ^2 検定および残差分析を行う。この検定により、有意差が認められれば、（4）②の仮の定量的な数値が、検証されることになる。

3.3.4 結果

3.3.4.1 重回帰分析の結果

研究課題（1）の解答を得るために、自由英作文の WS を従属変数とし、言語的特徴を表す 5 つの視点から抽出した 10 変数を説明変数候補とする重回帰分析を行った。重回帰分析はステップワイズ法を採用した。ステップワイズ法は統計的に最も予測に寄与する独立変数が順に投入され、適合度が最良の重回帰式を調べる際に使われるが（平井，2012）、誤差にも影響されやすく、結果を一般化しにくいという限界がある。表 3-11 はその結果を示し、Model 1 が従属変数である自由英作文の全体的評価の予測に役立つことも判明した（ $F=66.47$, $p<.001$ ）。自由英作文の WS の要因として抽出された説明変数は総語数、CLI、MTLD、Error-free T-unit 平均語数の 4 変数である。Error-free T-unit 平均語数に関しては p 値が 0.05 となり、有意傾向ではあるが、その数値から有意性に限りなく近いと判断し、説明変数に加えても差し支えないものと判断した。これらの 4 説明変数により $R^2=0.73$ のため、説明率は約 73% である。

説明変数間の関係に関しては、総語数と MTLD、総語数と Error-free T-unit 平均語数、MTLD と Error-free T-unit 平均語数のピアソンの相関係数がそれぞれ、0.24 ($p<.05$ 両側検定)、0.35 ($p<.01$ 両側検定)、0.39 ($p<.01$ 両側検定) のように低い正の相関関係があるが、他の変数間ではほとんど相関関係がみられない数値となっている（表 3-12 参照）。このように低い正の相関関係はあるものの、 $R^2=0.73$ に対し、調整済み $R^2=0.71$ であり、ほとんど数値が下がっていないことから、これらの説明変数間は互いに独立しているとみなしても差し支えないと考えられる。許容度や VIF の数値については総語数の許容度 (0.85)・VIF (1.17)、CLI 許容度 (0.94)・VIF (1.07)、MTLD の許容度 (0.83)・VIF (1.21)、Error-free T-unit 平均語数の許容度 (0.76)・VIF (1.31) となっており、許容度が 0.50 に満たない、もしくは、VIF が 2.00 を超える数値がなく、多重共線性にもなっていない。標準化係数 β 値から判断すると、総語数、CLI、MTLD、Error-free T-unit 平均語数はそれぞれ 0.74 : 0.19 : 0.12 : 0.12 の比率で自由英作文の WS に貢献することが判明した。

表 3-11 重回帰分析の結果

Model 1	非標準化係数		標準化係数		t 値	p	許容度	VIF
	R ²	B	標準誤差	β				
定数項		24.88	5.66		4.40	.00		
総語数	.64	.47	.04	.74	13.01	.00	.85	1.17
CLI	.04	1.91	.54	.19	3.57	.00	.94	1.07
MT LD	.14	.09	.04	.12	2.08	.04	.83	1.21
Error - free	.21	1.35	.68	.12	1.99	.05	.76	1.31
T-unit 平均語数								

R=.85 R²=.73 調整済み R²=.71

表 3-12 相関行列表

	(1)	(2)	(3)	(4)	(5)
(1) WS	—	.80**	.21*	.37**	.46**
(2) 総語数		—	-.04	.24*	.35**
(3) CLI			—	.17	.19
(4) MTL D				—	.39*
(5) Error-free T-unit 平均語数					—

**： p<.01 *： p<.05

3.3.4.2 説明変数の定量的な数値

研究課題 (2) の解答を得るために、4.3.3.4 の手続き (4) および (5) の結果を以下に示す。WS の得点の上位者 (100 点～130 点：60 名)、下位者 (100 点未満：46 名)^{注28)} とした。上位者 60 名と下位者 46 名の人数の偏りを生じたが、t 検定の際に 2 グループのサンプル数の偏りをできるだけ抑える必要性があり、目安は 1.5 倍程度としている (竹内・水本, 2012, p.63) ため、この偏りはその目安の範疇である。独立 2 群の平均値の差の検定の場合、それぞれの標本が正規分布であり、等分散性の 2 つの条件を満たさなければならない。この 4 変数は正規分布していたが (Kolmogorov-Smirnov の検定)、等分散性におい

注28) WS はグレード 6 (160 点)・グレード 5 (150 点～130 点)・グレード 4 (120 点～100 点)・グレード 3 (90 点～80 点)・グレード 2 (70 点～40 点)・グレード 1 (40 点未満) の 6 段階に分かれている。本研究の参加者はグレード 5 (130 点)～グレード 2 (50 点) に分布している。グレード 4 の説明に「課題に沿った話の展開が十分にできている。論理的に整理された文章が書けている。ごくまれにミスによって考えが伝わりにくいことがある」とあり、グレード 3 の説明は「話の展開はやや不十分だが、具体的な事例を含めて、ほぼ課題に沿った内容が書けている。時にミスによって考えが伝わりにくいことがある」と記されている。1 章 4 節の自由英作文の定義から、書き手の考えの伝わり方を重要と考え、その差異が顕著となるグレード 3 以下を下位者、グレード 4 以上を上位者と定義づけた。

ては、総語数について Levene 検定により等分散性ではないことが判明した（総語数： $F(1, 104)=12.15, p=0.00$ ）。従って、総語数の検定に関しては Welch の t 検定を行うこととした。

表 3-13 に各説明変数に関する WS 上位者および下位者の平均値および t 検定の結果を示す。総語数に関しては、Welch の t 検定の結果、WS の上位者と下位者の総語数の平均値の差は有意であった（両側検定： $t(99)=10.46, p=0.00$ ）。また、 t 検定の効果量に関しては、Hedges の g を使用することとし、効果量の目安は「0.20 (小)、0.50 (中)、0.80 (大)」(平井, 2012, p. 48) に従った。Hedges の効果量を算出した結果、 $g=1.92$ となり、効果が大きいことがわかった。CLI に関しては、 t 検定の結果、WS の上位者と下位者の CLI の平均値の差は有意であった（両側検定： $t(104)=2.08, p=0.04$ ）。Hedges の効果量を算出した結果、 $g=0.41$ となり、効果が小さいことがわかった。MTLD に関しては、 t 検定の結果、WS の上位者と下位者の MTLD の平均値の差は有意であった（両側検定： $t(104)=2.22, p=0.03$ ）。Hedges の効果量を算出した結果、 $g=0.44$ となり、効果が小さいことがわかった。Error-free T-unit 平均語数に関しては、 t 検定の結果、WS の上位者と下位者の Error-free T-unit 平均語数の平均値の差は有意であった（両側検定： $t(104)=4.26, p=0.00$ ）。Hedges の効果量を算出した結果、 $g=0.84$ となり、効果が大きいことがわかった。

表 3-13 総語数・CLI・MTLD・Error-free T-unit 平均語数に関する WS の上・下位者の平均値

言語的特徴を表す変数	WS 上位者		WS 下位者		2つの母平均の差の検定			
	$N=60$	$M(SD)$	$N=46$	$M(SD)$	t 値	自由度	有意確率 (両側)	効果量 g
総語数	116.63	(25.29)	75.30	(15.09)	10.46	98.63	.00	1.92
CLI	6.29	(1.93)	5.52	(1.82)	2.08	104.00	.04	0.41
MTLD	76.16	(26.72)	65.55	(20.87)	2.22	104.00	.03	0.44
Error-free T-unit 平均語数	8.35	(1.59)	7.05	(1.51)	4.26	104.00	.00	0.84

このように各変数の WS 上位者と WS 下位者の平均値には有意差が認められたため、仮の定義として、言語的特徴を表す変数の上位者の平均値以上（総語数：117 語以上；CLI：6.29 以上；MTLD：76.16 以上；Error-free T-unit 平均語数：8.35 以上）、上位者と下位者の平均値の間（総語数：116 語～76 語；CLI：6.28～5.53；MTLD：76.15～65.56；Error-free T-unit 平均語数：8.34～7.06）、下位者の平均値以下（総語数：75 語以下；CLI：5.52 以下；MTLD：65.55 以下；Error-free T-unit 平均語数：7.05 以下）に分けてみた。さらに、WS 上位者、WS 下位者の人数と各変数の上位者、中位者、下位者の人数の 2 条件×3 値でクロス集計し、 χ^2 検定を行い、有意差があるかどうかを検証した。

表 3-14 は WS の上・下位者と総語数の上・中・下位者のクロス集計表である。χ²検定の結果、人数の偏りは有意である (χ²(2)=50.25, p=0.00)。また、効果量として Cramer の V が 0.69 となった。Cramer の V に関する効果量の一般的基準「0.1 (効果量小)、0.3 (効果量中)、0.5 (効果量大)」からすると、効果量大と判断される (竹内・水本, 2012, p. 353)。残差分析の結果、総語数 75 語以下だと、WS の下位者は増え、総語数 117 語以上だと、WS の上位者が増えることも判明した。従って、仮の定義は検証された。

表 3-15 は WS の上・下位者と CLI の上・中・下位者のクロス集計表である。χ²検定の結果、人数の偏りは有意である (χ²(2)=7.30, p=0.03)。また、効果量として Cramer の V が 0.26 となったため、効果量中程度と判断される。残差分析の結果、CLI が 5.52 以下だと、WS の下位者は増え、CLI が 6.29 以上だと、WS の上位者が増えることも判明した。従って、CLI に関して、仮の定義は検証された。

表 3-16 は WS の上・下位者と MTLD の上・中・下位者のクロス集計表である。χ²検定の結果、人数の偏りは有意ではない (χ²(2)=0.65, p=0.72)。従って、MTLD に関して、仮の定義は検証されなかった。

表 3-17 は WS の上・下位者と Error-free T-unit 平均語数の上・中・下位者のクロス集計表である。χ²検定の結果、人数の偏りは有意である (χ²(2)=19.39, p=0.00)。また、効果量として Cramer の V が 0.43 となったため、効果量中程度と判断される。残差分析の結果、Error-free T-unit 平均語数が 7.05 以下だと、WS の下位者は増え、Error-free T-unit 平均語数が 8.35 以上だと、WS の上位者が増えることも判明した。従って、Error-free T-unit 平均語数に関して、仮の定義は検証された。

表 3-14 WS の上・下位者と総語数の上・中・下位者のクロス集計表

		総語数			合計
		上位者 117 語以上	中位者 116 語～76 語	下位者 75 語以下	
WS	度数	29	29	2	60
上位者	調整済み残差	5.5	0.5	-6.2	
WS	度数	0	20	26	46
下位者	調整済み残差	-5.5	-0.5	6.2	
合計	度数	29	49	28	106

表 3-15 WS の上・下位者との CLI の上・中・下位者のクロス集計表

		CLI			
		上位者	中位者	下位者	
		6.29 以上	6.28～5.53	5.52 以下	合計
WS	度数	34	3	23	60
上位者	調整済み残差	2.7	-0.3	-2.5	
WS	度数	14	3	29	46
下位者	調整済み残差	-2.7	0.3	2.5	
合計	度数	48	6	52	106

表 3-16 WS の上・下位者と MTLD の上・中・下位者のクロス集計表

		MTLD			
		上位者	中位者	下位者	
		76.16 以上	76.15～65.56	65.55 以下	合計
WS	度数	24	10	26	60
上位者					
WS	度数	15	8	23	46
下位者					
合計	度数	39	18	49	106

表 3-17 WS の上・下位者と EFT* の上・中・下位者のクロス集計表

		Error-free T-unit 平均語数			
		上位者	中位者	下位者	
		8.35 以上	8.34～7.06	7.05 以下	合計
WS	度数	33	15	12	60
上位者	調整済み残差	4.2	-0.9	-3.5	
WS	度数	7	15	24	46
下位者	調整済み残差	-4.2	0.9	3.5	
合計	度数	40	30	36	106

*: Error-free T-unit 平均語数

3.3.5 考察

研究課題(1)「CEFR-J A2 レベルの学習者が記述する自由英作文の言語的特徴として有効な説明変数は何か」については、重回帰分析の結果より、総語数、CLI、MTLD、Error-free T-unit 平均語数となった。これらの4変数はCEFR-J A2 レベルの学習者が記述する自由英

作文のルーブリックにおける評価の観点になりうる可能性がある。4変数の内で、総語数の標準化係数 β 値が0.74、CLIのそれが0.19、MTLDおよびError-free T-unit平均語数が0.12であり、このモデルにおける総語数の自由英作文に対する評価への貢献度の大きさが明確である。つまり、 β 値の比率はほぼ、総語数：CLI：MTLD：Error-free T-unit平均語数=6：1.5：1：1となる。

モデルの説明率も73%程度となり、先行研究のFerris (1994)の50%程度、杉浦 (2008)の44%程度、水本 (2008)の60%程度と比較しても、より有効なモデルであることを示している。抽出された説明変数はFluencyの視点から総語数、文の視点からError-free T-unit平均語数、語彙の視点からMTLD (語彙の多様性)、リーダビリティの視点からCLIが抽出されていることになる。

先行研究の中でも水本 (2008)は調査対象の英語力が本研究と最も近似しているため、両者を比較して考察する。本研究では上述した4変数が説明変数として抽出されたが、水本 (2008)では総語数、FKG、および平均語長が説明変数として抽出されている。両研究で説明変数候補として共通するのはfluency、語彙、リーダビリティの3視点からの変数である。Fluencyからは総語数が共に説明変数として抽出され、リーダビリティも指標は異なるものの、説明変数が抽出されている。総語数に関しては、標準化係数 β 値はともに0.7程度で、抽出された説明変数の中ではいずれの研究においても最も大きな値になっている。語彙の視点からは水本 (2008)では平均語長が抽出されているが、本研究では抽出されなかった。また、語彙の多様性が本研究では抽出 (MTLD)されたが、水本 (2008)では語彙の多様性を示す3指標がいずれも抽出されていない。これらの相違は水本 (2008)の調査対象の英語力がTOEIC 300点前半と推定される一方、本研究ではCEFR-J A2レベルの英語力の範囲に近い242点～548点の対象者としており、英語力にかなりの相違があることが一因とも考えられる。また、大学生と高校生という学習者の年齢の相違によって生じているかもしれない。この点については今後の検証が必要である。

談話の視点からの結束性に関して、橋内 (1995)の指摘するように初学者の自由英作文における結束性は重要であるが、本研究において抽出されなかったのは、ここでは、TOEIC 400点を超え、英検2級程度の学習者を含んでいるため、橋内 (1995)の初学者の範疇を超える学習者を含んでいるためと考えられる。

これらの4変数をもとにCEFR-J A2レベルの学習者に自由英作文を指導する際の評価に関する示唆を以下にまとめる。

- (1) 総語数の多い作文は高評価となる可能性が高くなる。
- (2) Error-free T-unitの平均語数の数値が高いほど、書き手の考えや意見を読み手に伝達するのに支障をきたすようなエラー^{注29)}が少なく、統語的に熟達した文と評価され

注29) このエラーは “an error in the use of a major element of sentence structure, which makes a sentence or utterance difficult or impossible to understand.” (Richards

る可能性が高くなる。

(3) MTLD の数値が高く、多様な語彙を使用した作文は高評価を受ける可能性が高くなる。

但し、CEFR-J A2 レベルの学習者の語彙の使い方の特性から、MTLD の数値で判断するには注意を要する（研究課題（2）の考察参照）

(4) CLI の数値の高いリーダビリティの優れた作文は高評価を受ける可能性が高くなる。研究課題(2)「自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値は何か」については MTLD を除いて、「3.3.4.2 説明変数の定量的な数値」の自由英作文の習熟度による言語的特徴を表す説明変数に関する仮の定義の数値が検証された。従って、解答として総語数「上位者：117 語以上；中位者：116 語～76 語；下位者：75 語以下」、CLI「上位者：6.29 以上；中位者：6.28～5.53；下位者：5.52 以下」、Error-free T-unit 平均語数「上位者：8.35 以上；中位者：8.34～7.06；下位者：7.05 以下」という数値を提示する。このような数値化は自由英作文のルーブリックの3段階の評価基準（到達度合い）を決定する基礎データとなる。表 3-18 に CEFR-J A2 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値を示した。

MTLD に関して仮の定義が支持されなかったのは、本研究の調査参加者における語彙使用はライティング力のいずれの習熟度においても『CEFR-J Wordlist Version 1.0』（東京外国語大学投野由紀夫研究室）の A1 レベルに 90%程度が集中^{注30)}していたため、語彙の多様性についてはライティング力の習熟度による差異がさほど大きくなかったことに原因があると考えられる。

表 3-18 CEFR-J A2 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値

	上位者	中位者	下位者
総語数	117 以上	116～76	75 以下
CLI	6.29 以上	6.28～5.53	5.52 以下
Error-free T-unit 平均語数	8.35 以上	8.34～7.06	7.05 以下

3.3.6 結論

3 章 3 節では、CEFR-J A2 レベルの自由英作文のルーブリックにおける評価の観点と評価基準（到達度合い）の基礎データを探求した。評価の観点は自由英作文の言語的特徴を重回帰分析による説明変数である特定要因の評価指標により抽出した。評価基準（到達度合

et al. ,1992, p. 157) という定義に従い、グローバルエラーを対象とした。従って、3 人称単数現在の s や冠詞のエラー、単純なスペルミスのような局所的エラーは本研究ではエラーとして算出しないこととした。

注30) 3 章 2 節参照

い) は説明変数となった特定要因の評価指標に関して、自由英作文の全体的評価を基に、統計的に検証した定量的数値により 3 段階に区分した。以下に、これらの分析により得られた知見を示す。

- (1) CEFR-J A2 レベルの自由英作文の言語的特徴は総語数、CLI、MTLD、Error-free T-unit 平均語数の 4 つの特定要因の評価指標となった。これらの 4 変数は自由英作文のルーブリックにおける評価の観点になる可能性がある。
 - (2) (1) の抽出された言語的特徴の中で、語彙の多様性を表す指標である MTLD を除く、3 指標に関して、統計的に検証した結果、3 段階の定量的な数値を得ることができた。すなわち、CEFR-J A2 レベルの自由英作文のルーブリックにおいて、総語数、CLI、Error-free T-unit 平均語数に関しては、3 段階の評価基準（到達度合い）を有する評価の観点とすることが可能であることが明らかとなった。
 - (3) 評価の観点となる総語数、CLI、Error-free T-unit 平均語数の重回帰分析の標準化係数 β 値の比率は 6:1.5:1 となった。
 - (4) 評価の観点となる総語数、CLI、Error-free T-unit 平均語数のそれぞれの評価基準（到達度合い）の定量的数値は表 3-18 のようになった。
- これらの結果を基に、CEFR-J A2 レベルの自由英作文のルーブリックをデザインする。

3.4 まとめ

3 章では、CEFR-J A2 レベルの自由英作文についてリーダビリティと語彙の豊かさの視点から分析した。その結果を踏まえて、重回帰分析により CEFR-J A2 レベルの自由英作文の言語的特徴を抽出し、ルーブリックにおける評価の観点を探ることとした。

リーダビリティの視点から分析した結果として、CEFR-J A2 レベルの自由英作文を読み、理解するにはアメリカの小学校 3 年生から 6 年生程度の教育レベルが必要なこと、また、リーダビリティの視点から分析する際に使用したリーダビリティの指標である CLI は CEFR-J A2 レベルの自由英作文の言語的特徴を表わす特定要因の評価指標になりうる可能性があることが判明した。

語彙の豊かさ（多様性と広範さ）の視点から分析した結果として、語彙の広範さよりも多様性の方が CEFR-J A2 レベルの自由英作文の言語的特徴になりうる可能性が高いことが判明した。多様性と広範さの関係については、両者の重なる部分が少ないことが明らかとなった。これは CEFR-J A2 レベルという初級レベルの英語力のメンタルレキシコンの未発達に原因があると考えられる。CEFR-J A2 レベル（TOEIC 225 点～550 点）内の習熟度では TOEIC 400 点を閾値として、多様性に関して、語彙の使用が区別されることも明らかとなった。また、CEFR-J A2 レベルの自由英作文で使用される語彙は『CEFR-J Wordlist Version 1.0』（東京外国語大学投野由紀夫研究室）の A1（小学校～中学校 2 年程度）という高頻度語彙に 90% 程度が集中していることも明らかとなった。

CEFR-J A2 レベルの自由英作文に関して、重回帰分析により言語的特徴を抽出した結果、

総語数、CLI（リーダビリティの指標）、MTLD（語彙の多様性の指標）、Error-free T-unit 平均語数の4つの特定要因の評価指標を得られた。これらの指標は本研究におけるルーブリックの評価の観点となる可能性がある。この4指標に関して、ルーブリックの評価基準（到達度合い）の作成ため、統計的に有意になるように3段階の定量的な数値の算出を試みた。結果として、MTLDについては、統計的に有意な3段階の定量的な数値を得られず、ルーブリックの評価の観点として利用することができないことが明らかとなった。他の3指標については統計的に有意な3段階の定量的な数値を得られたため、ルーブリックの3段階の評価基準（到達度合い）を有する評価の観点として利用できることが判明した。従って、CEFR-J A2レベルの自由英作文のルーブリックについては総語数、CLI、Error-free T-unit 平均語数の3つの特定要因の評価指標を評価の観点とすることとなった。

4. CEFR-J B2 および C1 レベルの自由英作文の評価の観点の提案

4.1 CEFR-J B2 および C1 レベルの自由英作文のリーダビリティによる分析

4.1.1 はじめに

4章1節では、CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文をリーダビリティの視点から分析する。分析手法は3章1節に準ずることとする。

リーダビリティの指標が自由英作文のようなライティングの分析に活用されてこなかったことは、3章1節でも記述した。そのため、CEFR-J B2 および C1 レベルのような上級の外国人英語学習者によって記述された自由英作文のリーダビリティがどの程度になるか、全く把握されていないのが現状である。3章1節の結果として、CEFR-J A2 レベルの自由英作文のリーダビリティとして、アメリカの小学校3年生～6年生程度を表わす数値を得られた。常識的に考えれば、CEFR-J B2 および C1 レベルの自由英作文のリーダビリティは、CEFR-J A2 レベルのリーダビリティよりも高くなることは確実であろうが、どの程度の数値になるのかは判然としない。また、3章1節の結果として、CEFR-J A2 レベルの自由英作文の全体的評価とリーダビリティの指標には弱い正の相関関係が認められ、リーダビリティの指標はCEFR-J A2 レベルの自由英作文の言語的特徴になりうる可能性が示唆された。実際、3章3節の結果、リーダビリティの指標であるCLIはCEFR-J A2 レベルの言語的特徴を表わすことも実証された。

このような結果を踏まえ、本節ではCEFR-J B2 および C1 レベルの自由英作文をリーダビリティの視点から分析し、本研究の主目的であるループリックの評価の観点としての自由英作文の言語的特徴としてリーダビリティの指標が有用であるかどうか、もし、有用であるならば、どの指標を用いるのが適切であるかを調査する。

4.1.2 研究の背景

研究の背景としては、3章1節の研究の背景に準拠する。但し、分析前の段階では、CEFR-J B2 および C1 レベルの自由英作文のリーダビリティがどの程度となるか判然としないため、染谷(2009:10-11)で提示された英文サンプルのリーダビリティが目安になると考え、それを以下に示すこととした。そこでは、次のような13本の英文サンプルのリーダビリティが提供されている。

- (1) 筆者(染谷教授)の勤務校(青山学院大学)で実際に使用している英文購読用教材から抜粋した英文エッセイ4本:表5-1のSample 1~4
- (2) リーダビリティ評価用に筆者(染谷教授)が授業で使用しているビジネスレターの例文7本:表5-1のBL-07, BL-08
- (3) レファレンス用のテキストとして米国大統領の就任演説を2本

CEFR-J A2 レベルの自由英作文の自由英作文のリーダビリティの指標が3.37~6.00(表3-2参照)を示したことから、CEFR-J B2 および C1 レベルの自由英作文は、その英語力の相違から少なくともリーダビリティの指標が6.0を超えるのは確実と考えられる。表4-1

のエッセイやビジネスレターの難易度の普通程度に留まるか、やや難もしくは難のレベル、さらにそれ以上のレベルにまで到達するかを見極める必要がある。

表 4-1 サンプルテキストの分析結果（染谷（2009：11）から引用）

							テキストタイプと主観的評価による
N	TEXT ID.	FKG	ARI	CLI	Ave.	STD	難易度 (RF=Reference Text)
1	Sample 1	8	7.4	7.9	7.8	0.32	エッセイ：購読用教材からの抜粋（易）
2	Sample 2	14	13.6	13.6	13.7	0.23	エッセイ：同上（やや難）
3	Sample 3	10	10.8	9.6	9.6	0.61	エッセイ：同上（普通）
4	Sample 4	14	15.0	14.7	14.6	0.51	エッセイ：同上（難）
5	Speech 1	11	11.1	11.2	11.1	0.10	米国大統領就任演説 (Bush) 2006 (RF)
6	Speech 2	10	10.3	10.2	10.2	0.15	米国大統領就任演説 (Clinton) 2006 (RF)
7	BL-07	10	9.3	9.9	9.7	0.38	ビジネスレター（普通）
8	BL-08	8	7.0	8.8	7.9	0.90	ビジネスレター（易）
9	BL-13	10	9.2	8.0	9.1	1.01	ビジネスレター（普通）
10	BL-19	11	9.6	11.9	10.8	1.16	ビジネスレター（普通）
11	BL-23	14	15.7	15.6	15.1	0.95	ビジネスレター（やや難）
12	BL-30	13	13.1	12.1	12.7	0.55	ビジネスレター（やや難）
13	BL-60	5	4.3	5.0	4.8	0.40	ビジネスレター（易）

表 4-1 の結果を踏まえて、CEFR-J B2 および C1 レベルの自由英作文をリーダビリティの視点から分析していくこととする。

4.1.3 研究の目的

本節では、CEFR-J B2 および C1 レベルの学習者が記述する自由英作文はリーダビリティの視点から分析し、そのテキストを読むために最低限必要とされる教育レベルはアメリカの学年で、どのくらいのレベルになるのかを明らかにし、そのリーダビリティの指標が本研究の主目的であるルーブリックの評価の観点、言い換えれば、自由英作文の言語的特徴の説明変数候補として有用かどうか、もし、有用と考えられるならば、どの指標を用いるのが妥当なのかを調査することを目的とする。本研究で取り扱うリーダビリティの指標として、最も一般的な FKG、ARI、CLI を用いることとする。

以下のこれらの目的を達成するために 2 つの研究課題を設定する。

- (1) CEFR-J B2 および C1 レベルの学習者が記述した自由英作文のテキストはリーダビリティの指標である FKG、ARI、CLI の平均値はどうなるのか。
- (2) リーダビリティの指標である FKG、ARI、CLI の中で、CEFR-J B2 および C1 レベルの学

習者が記述した自由英作文を評価するのに、どの指標が最も妥当性が高いか。

4.1.4 方法

4.1.4.1 CEFR-J の B2 および C1 レベルの調査参加者

本研究にあたり、英語のプロフェッショナルを養成する A 英語学校に協力を受諾してもらうことができた。同校では英語技能検定 1 級のような上級資格の取得を希望する受講生が多数在籍している。そのような受講生（修了生も含む）の中で、CEFR-J B2 レベル (TOEIC 785 点) 以上の英語力を有する日本人上級英語学習者 65 名（男性 25 名、女性 40 名）を調査参加者とした。CEFR では B2 レベルは自立した言語使用者の上位、C1 は熟達した言語使用者（吉島・大橋他，2004）と位置づけられているが、TOEIC 785 点以上の英語力を有する日本人学習者は一般に上級学習者と判断して差し支えないのではないかと考える。CEFR-J B2 レベル (TOEIC 785 点～945 点) が 34 名、CEFR-J C1 レベル (TOEIC 945 点～990 点) が 31 名である。年齢構成は 10 代が 2 名、20 代が 4 名、30 代が 23 名、40 代が 21 名、50 代が 10 名、60 代が 5 名である。表 4-2 は調査参加者の英語力の記述統計である。TOEIC での平均点が 907 点、最小値は 785 点、最大値は 990 点である。

表 4-2 調査参加者の英語力の記述統計 (N = 65)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
TOEIC R	495	448.92	43.23	345	495	150	-0.57	-1.05
TOEIC L	495	458.00	25.60	395	495	100	-0.56	-0.78
合計	990	907.15	66.06	785	990	205	-0.51	-1.21

TOEIC R: リーディングセクション TOEIC L: リスニングセクション

4.1.4.2 ライティングテスト

‘Agree or disagree : when people move to another country, people should follow the customs of the new country.’ という課題で調査協力者に自由英作文を記述してもらい、これを本研究におけるライティングテストとした。その自由英作文を指導歴 10 年以上の 2 人の評価者によって、印象度による全体的評価で採点した。

採点は TOEFL の Independent 型 Writing (本研究の自由英作文に相当) の得点 0 点～5 点までの 6 段階に準拠したが、各段階の中間点 (0.5 点を除く、1.5、2.5、3.5、4.5) を加えた 10 段階とした。評価者間信頼性を調べてみたところ、カッパ係数は 0.77 であった。この 2 人の評価者の得点の平均値をこの自由英作文の全体的評価 (holistic evaluation: HE) の得点とした。表 4-3 に記述統計を示した。

表 4-3 自由英作文の全体的評価 (N = 65)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
全体的評価 (HE)	5.00	4.03	0.65	2.00	5.00	3.00	-0.26	0.04

4.1.4.3 リーダビリティの指標

3章1節の「CEFR-J A2 レベルの自由英作文のリーダビリティによる分析」で使用したFKG、ARI、CLIを用いることとする。

4.1.4.4 手続き

(1) 自由英作文のテキストデータの収集

①CEFR-J B2 および C1 レベルの日本人上級英語学習者に自由英作文を記述してもらう。

(2) 自由英作文の全体的評価

①CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文を ETS によって作成された ‘Independent WRITING Rubrics’ を基に、2 人の評価者による印象度による評価である全体的評価によって採点する。

②2 人の評価者は、自由英作文の採点に 10 年以上の経験のある A 英語学校の講師である。

③採点は ‘Independent WRITING Rubrics’ では 0 点～5 点までの 6 段階であるが、各段階の中間点 (0.5 点を除く、1.5、2.5、3.5、4.5) を加えた 10 段階とする。

(3) リーダビリティの指標である FKG、ARI、CLI のスコアの算出

①FKG のスコアは MS-Word のスペルチェックと文章校正機能により算出した。

②ARI と CLI のスコアは前青山学院大学染谷泰正教授 (現関西大学教授) の作成された Word Level Checker (英文語彙難易度解析プログラム) により算出する (http://someya-net.com/wlc/index_J.html)。

(4) FKG、ARI、CLI に関する記述統計量の算出

(5) 全体的評価と FKG、ARI、CLI の相関関係の調査

4.1.5 結果と考察

表 4-4 はリーダビリティの指標に関する記述統計である。この結果から、平均値は CLI が最も高い 11.45、次に ARI の 11.18、最も低いのが FKG の 9.81 となった。この数値はアメリカの学年を示しているので、アメリカの学校制度における 10 年生レベルから 11 年生程度のリーダビリティということになる。

FKG は “authentic” な英文の場合、5.5 から 16.3 の範囲内に収まるとされるが、CEFR-J A2 レベルの自由英作文のテキストの場合、そのスコアは 4.22 であり、“authentic” な英文の範疇に入らないことを示したが、CEFR-J B2 および C1 レベルの自由英作文では 9.81 となり、“authentic” な英文の範疇に入ることとなった。

Taylor (2008) では、ARIはFKGやCLIよりも高めに算出されると指摘されている。FKGの平均値よりは、平均値で1.37高くなっているが、CLIよりは平均値で0.27低くなっている。しかしながら、ARIとCLIの差は非常に僅かなものであり、この程度の差は表5-1にもみられる。従って、CEFR-J A2レベルでARIが最も低い数値となったのとは状況は異なると思われるべきである。CEFR-J B2レベル以上になると、非母語話者といえども母語話者によって記述された表4-1のテキストとも同様な3指標間の差となるようである。

表 4-4 リーダビリティの指標に関する記述統計 (N = 65)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
FKG	9.81	1.36	4.90	15.00	10.10	0.43	0.04
ARI	11.18	2.38	6.50	17.50	11.00	0.42	0.30
CLI	11.45	1.65	7.90	17.10	9.20	0.41	1.27

表4-5は自由英作文の全体的評価とFKG、ARI、CLIの相関行列表である。自由英作文の全体的評価 (HE) とFKG、ARI、CLIのピアソンの相関係数は0.44、0.42、0.47であり、中程度の正の相関関係があるという結果となった。中程度の相関関係があるため、これらのリーダビリティの指標が、自由英作文の言語的特徴となる可能性は十分あると言えよう。

また、FKG、ARI、CLIの3指標間のピアソンの相関係数は、FKGとARIが0.93、ARIとCLIが0.79、CLIとFKGが0.71のように高い正の相関関係があることが判明した。これは、染谷(2009)における英文サンプルにおけるFKG、ARI、CLI間のピアソンの相関係数0.9を超える強い正の相関関係があるという結果とはやや異なることとなった。

表 4-5 自由英作文の全体的評価と FKG、ARI、CLI の相関行列表

	(1)	(2)	(3)	(4)
(1) HE (自由英作文の全体的評価)	—	.44*	.42**	.47**
(2) FKG		—	.93**	.71**
(3) ARI			—	.79**
(4) CLI				—

**： p<.01

このような結果から、以下のように考察した。研究課題(1)の「CEFR-J の B2 および C1 レベルの学習者が記述した自由英作文のテキストはリーダビリティの指標である FKG、ARI、CLI の平均値はどうなるのか」の解答は、FKG、ARI、CLI がそれぞれ 9.81、11.18、11.45 となり、CEFR-J B2 および C1 レベルの自由英作文を読むためには、アメリカの学校制度に

における 10 年生から 11 年生程度の最低限の教育レベルが必要なことが判明した。これは、表 4-1 のエッセイやビジネスレターの普通程度の難易度のテキストや米国大統領就任演説 (Bush 大統領、Clinton 大統領) に相当する。リーダビリティの各指標の平均値が 10~11 で、普通程度の難易度のテキストであるため、CEFR-J B2 および C1 レベルの学習者が記述した自由英作文は、アメリカ人のような英語母語話者の大人にとっては読みやすいテキストであることを示していると言えよう。

研究課題(2)の「リーダビリティの指標である FKG、ARI、CLI の中で、CEFR-J B2 および C1 レベルの学習者が記述した自由英作文を評価するのに、どの指標が最も妥当性が高いか」の解答は以下のように考察した。表 4-5 に示したように、FKG、ARI、CLI のいずれの指標も、自由英作文の全体的評価とは中程度の正の相関関係がある。このことは、これらのリーダビリティの指標が自由英作文の言語的特徴を表す指標として有用になる可能性を十分に示唆している。

自由英作文の言語的特徴を分析する際に、重回帰分析を使用するが、多重共線性を避けるため、説明変数候補間のピアソンの相関係数 0.8 以上としないという前提がある(平井, 2012)。そのため、FKG と ARI はピアソンの相関係数が 0.93 となり、0.8 を超えており、両者を説明変数候補とすることは控える必要がある。ARI と CLI はピアソンの相関係数が 0.79、CLI と FKG は 0.71 であり、0.80 以上とはなっていない。各指標の平均値として、ARI と CLI はそれぞれ 11.18 と 11.45 と近似値となっているが、FKG は 9.81 であり、ARI や CLI と比較してやや低い数値となっている。これらの 3 指標間のピアソンの相関係数と平均値を考慮して、FKG を除外し、ARI と CLI を重回帰分析の説明変数候補とすることとした。従って、研究課題(2)の解答として、CEFR-J B2 および C1 レベルの学習者が記述した自由英作文を評価するのに、ARI と CLI が最も妥当であると判断した。

4.1.6 結論

4 章 1 節では、CEFR-J B2 および C1 レベルの自由英作文をテキストの読みやすさおよび理解しやすさと定義したリーダビリティの指標である FKG、ARI、CLI により分析した。従来、リーダビリティはリーディングの研究に使用されることが一般的で、自由英作文のようなライティングにリーダビリティが活用された研究はあまり行われてこなかった。そのため、CEFR-J B2 および C1 レベルのような上級英語学習者によって記述された自由英作文とリーダビリティの関係は判明していなかったが、以下の事項が明らかとなった。

- (1) CEFR-J B2 および C1 レベルの自由英作文を読むためには、アメリカの学校制度における 10 年生から 11 年生程度の最低限の教育レベルが必要なこと。
- (2) リーダビリティが CEFR-J B2 および C1 の自由英作文の言語的特徴を表す指標として有用になる可能性があること。
- (3) CEFR-J B2 および C1 の自由英作文を評価するリーダビリティの指標は ARI と CLI が妥当であること。

この結論に基づき、2章5節3項の「CEFR-JのB2およびC1の自由英作文の評価の観点となる特定要因の評価指標候補」のリーダビリティの視点の説明変数候補から、FKGを除外することとした。

4.2 CEFR-J B2 および C1 レベルの自由英作文の語彙の豊かさの観点による分析

4.2.1 はじめに

4章2節においては、CEFR-J B2 および C1 レベルの上級英語学習者によって記述される自由英作文で使われる語彙の豊かさを多様性と広範さという観点から調査する。このレベルの学習者は、日本では一般的に上級者とみなされる人達である。海外の大学や大学院で学んだり、国際化著しいビジネス等の第一線で活躍するに相応しい英語力でもある。このような人達にとって、自らの考えを発信する自由英作文のようなライティング力の向上は学術の分野においても、また、ビジネスにおいても極めて重要である。

近年、日本人の学習者によって記述された自由英作文の語彙の研究（小島，2013；石川，2005，水本，2008；平林，2015 など）がなされてきている。しかしながら、CEFR-J B2 レベル以上の上級英語学習者によって記述された自由英作文において使用される語彙に関しては、小島（2013）において CEFR-J B2 レベルを対象として研究が行われているものの、CEFR-J C1 レベルに関しては、ほとんど研究がなされていないのが現状である。従って、CEFR-J B2 レベルおよび C1 レベルのような上級の学習者によって記述される自由英作文において使用される語彙に関して、必ずしも十分な知見が得られているとは言い難い。

3章2節で記述した研究の背景や CEFR-J A2 レベルの自由英作文の語彙の豊かさの分析結果から、語彙の豊かさを表わす指標が言語的特徴となりうることを示唆している。本節においても CEFR-J B2 および C1 レベルの自由英作文の語彙の豊かさを表わす指標が自由英作文の言語的特徴を表わすことが示唆されれば、それはルーブリックの評価の観点となりうる可能性があることを示している。また、CEFR-J B2 レベルおよび C1 レベルの学習者によって記述された自由英作文の語彙の豊かさの分析から新たな知見が得られれば、自由英作文の評価基準であるルーブリックの基礎データとなりうると思われる。

4.2.2 研究の背景

本節の研究の背景は3章2節の研究の背景に準拠するものとする。但し、以下の事項を付加する。

Laufer（1994）や Laufer & Nation（1995）の推奨する Lexical Frequency Profile（以下、LFP とする）に関して、海外の研究では使用されている（Muncie, 2002; Lemmouh, 2008; Nadarajan, 2011 など）が、日本の高校生や大学生を対象とした研究ではあまり使用されないようである。この理由として、葉田野・馬場（2006）は平均的な日本人大学生の英作文では General Service List (West, 1953) の高頻度語彙 1000 語の使用が多いため、LFP による分析が難しいことを指摘している。また、望月他（2003）では、LFP は作文の長さが 400 語以上の比較的長めの作文でなければ処理できないという問題点も指摘している。このため、CEFR-J A2 レベルの自由英作文の語彙の豊かさを調査した3章2節においても、400 語以上の自由英作文を対象とした 106 個の中には1つもなく、LFP は使用することができなかった。また、上級英語学習者を対象とする本節においても、研究対象とした 65 個の

自由英作文の総語数の平均は 229 語であり、358 語が最大の総語数であったため、LFP の使用は控えることとした。

本節でも小島 (2013) で使われた多様性と広範さに関する指標の使用を検討する。また、小島 (2013) で使われた語彙の豊かさに関する指標を用いた 3 章 2 節では、CEFR-J A2 レベル (TOEIC 225 点～550 点) の英語力の学習者によって記述された自由英作文においても語彙の豊かさは、その英文の質に一定の影響を及ぼすことが示唆された。豊かさの中でも、多様性の方が広範さよりも英文の質に反映していることも明らかとなった。また、多様性と広範さの指標間の相関関係が低いことから、両者の重なる部分は少ないのではないかと考えられる。さらに、CEFR-J A2 レベル内での英語の習熟度が高くなると、多様性はより明確になるが、広範さに関しては習熟度の影響はあまりみられないことを明らかにしている。

これらの指標に加え、3 章 2 節と同様に、『CEFR-J Wordlist Version 1.0』(東京外国語大学投野由紀夫研究室) という語彙リストを本研究における語彙の豊かさの調査に活用する。3 章 2 節では、『CEFR-J Wordlist Version 1.0』を基に語彙レベルを調査した結果、CEFR-J A2 レベルの自由英作文では、A1 の高頻度語彙に 90% 程度集中していることが明らかとなっている。本節でも、この語彙リストを使用して、上級英語学習者によって記述された自由英作文において、A1 のような高頻度語彙や B2 のような低頻度語彙の使用割合などを調査する。

4.2.3 研究の目的

本節では、CEFR-J B2 レベルおよび C1 レベルの日本人上級英語学習者が記述する自由英作文の語彙の豊かさを多様性と広範さの観点から分析し、その豊かさが自由英作文の質に影響するかどうか、また、多様性と広範さにおいてどのような使われ方をしているかを明らかにすることを目的とする。この目的のために、以下の研究課題を設定する。

- (1) 上級英語学習者が記述した自由英作文の全体的評価とそこで使用された語彙の豊かさ (多様性と広範さ) には相関関係があるのか。
- (2) 上級英語学習者が記述した自由英作文で使用された語彙の多様性と広範さには相関関係があるのか。
- (3) 英語の習熟度 (2 群:CEFR-J B2 レベルおよび CEFR-J C1 レベル) が異なると、語彙の豊かさには差異があるのか。
- (4) 上級英語学習者が記述した自由英作文では、CEFR-J の語彙リスト『CEFR-J Wordlist Version 1.0』のどのレベルの語彙を使用しているのか。
- (5) 『CEFR-J Wordlist Version 1.0』では、英語の習熟度 (CEFR-J B2 レベルおよび CEFR-J C1 レベル) により、語彙レベルの使われ方には差があるのか。

4.2.4 研究方法

4.2.4.1 CEFR-J の B2 および C1 レベル調査参加者

4章2節におけるCEFR-JのB2およびC1レベル調査参加者は4章1節と同一である(表4-2参照)。

4.2.4.2 ライティングテスト

4章2節におけるライティングテストは4章1節と同一である(表4-3参照)。

4.2.4.3 語彙の豊かさの指標の記述統計

語彙の豊かさを測定する指標は、当初、小島(2013)で使われた多様性の指標であるGI、D、MTLD、広範さに関しては、P_Lex、Sの5つを使用することを検討していた(語彙の豊かさに関する指標に関しては3章2節参照)。

しかしながら、多様性の指標に関しては、3章2節において記述したように、GIはTTRの総語数の影響を受けるという欠点がある程度緩和するものの必ずしも十分とは言えないと判断し、総語数の影響を受けにくくするために開発されたDとMTLDを採用することとした。

また、広範さに関する指標に関しては、指標Sが安定した数値となるためには、200語以上の英文テキストを推奨しており、本研究における65個の自由英作文の内、19個の総語数が200語を下回るために、Sを使用することを控えることとした。

従って、本研究では多様性に関する指標としてDとMTLD、広範さに関する指標としてP_Lexを使用することとした。以下、表4-6に各指標に関する記述統計を示す。

表4-6 語彙の豊かさの指標に関する記述統計

N=65	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
D	57.77	17.05	29.51	99.99	70.46	0.74	0.08
MTLD	78.71	25.76	39.59	204.55	164.96	2.04	7.84
MTLD#	1.88	0.13	1.60	2.31	0.71	0.44	1.24
P_Lex	1.42	0.58	0.38	2.73	2.35	0.46	-0.36

MTLD#: 数値変換後のMTLD MTLDに関しては正規性に関して問題があったため、数値変換を行った。

4.2.4.4 手続き

(1) 自由英作文のテキストデータの収集

日本人のCEFR-J B2レベルおよびC1レベルの上級英語学習者65名に自由英作文を記述してもらい、その英文テキストをMS-Wordにより、デジタルデータ化した。

- (2) 自由英作文を Educational Testing Service (以下、ETS とする) によって作成された ‘Independent WRITING Rubrics’ を基に全体的評価で採点した。
- (3) 語彙の多様性の指標である D、MTLD の値を算出
D は free software である D_Tools (<http://www.lognostics.co.uk/tools/index.htm> からダウンロード) を基に値を算出した。
MTLD は Web 上の (<http://gnutiez.de/wp/apps/mtld/>) プログラムによって値を算出した。
- (4) 語彙の広範さの指標である P_Lex の値の算出
P_Lex は分析ツール (<http://www.lognostics.co.uk/tools/> からダウンロード) を使用して値を算出した。
- (5) 自由英作文の全体的評価と各指標 (D、MTLD、P_Lex、) の相関関係の調査
- (6) 各指標 (D、MTLD、P_Lex) 間の相関関係の調査
- (7) 各指標に関して、英語の習熟度の異なる 2 群 (CEFR-J B2 レベルと CEFR-J C1 レベル) において、その平均値に有意差があるのかどうか検証する。
- (8) 上級英語学習者によって記述された自由英作文の発表語彙が、『CEFR-J Wordlist Version 1.0』のどのレベルの語彙を使用しているか調査する。
- (9) 上級英語学習者によって記述された自由英作文の発表語彙が、『CEFR-J Wordlist Version 1.0』において、英語の習熟度の異なる 2 群 (CEFR-J B2 レベルと CEFR-J C1 レベル) において、語彙レベルの使われ方に差があるのか調査する。

4.2.5 結果

4.2.5.1 自由英作文の全体的評価と語彙の豊かさに関する指標の関係と各指標間の関係

表 4-7 は自由英作文の全体的評価と語彙の豊かさに関する指標の相関行列表である。語彙の多様性を表す D、MTLD# と自由英作文の全体的評価とのピアソンの相関係数はそれぞれ $r = .11$ 、 $r = .12$ となり、ほとんど相関関係がないことが明らかとなった。一方、語彙の広範さを表す P_Lex と自由英作文の全体的評価とのピアソンの相関係数は $r = .46$ となり、中程度の正の相関関係があることが判明した。

また、語彙の多様性を表す指標 (D、MTLD#) と語彙の広範さを表す指標 (P_Lex) の間にはいずれも中程度の正の相関関係があることが判明した。

表 4-7 自由英作文の全体的評価と語彙の豊かさに関する指標の相関行列表

	HE	D	MTLD#	P_Lex
HE	--	.11	.12	.46**
D		--	.78**	.51**
MTLD#			--	.40**
P_Lex				--

HE: 自由英作文の全体的評価 **: $p < .01$

4.2.5.2 英語の習熟度による発表語彙の豊かさの指標の比較

表 4-8 は習熟度による語彙の豊かさの指標のスコアを比較したものである。多様性を表す指標 D および MTLD#、広範さを表す指標 P_Lex に関して、英語力の習熟度の相違 (CEFR-J B2 レベルと CEFR-J C1 レベル) により、その平均値の差を対応なしの t 検定により分析した。また、等分散性を確認するために Levene の検定を行った結果、P_Lex に関しては等分散性がないことが判明した。 t 検定の結果、D に関しては、習熟度の差により有意差はみられなかった ($t(63) = 1.68, p = .1$)。MTLD# に関しても習熟度の差により有意差はみられなかった ($t(63) = 1.40, p > .1$)。従って、語彙の多様性に関しては習熟度の差による有意差はみられないと判断した。一方、P_Lex に関しては、習熟度により有意差がみられた ($t(51) = 4.16, p < .01$)。その効果量 Δ は 0.89 となり、効果が大きいこともわかった。従って、語彙の広範さに関しては習熟度の差により有意差がみられると判断した。

表 4-8 習熟度による語彙の豊かさに関する指標のスコアの相違

発表語彙の豊かさ を表す指標	C1 レベル		B2 レベル		2つの母平均の差の検定			
	N=31	M(SD)	N=34	M(SD)	t 値	自由度	有意確率 (両側)	効果量 Δ
D		61.44 (19.38)		54.43 (14.09)	1.68	63	.10	
MTLD#		1.90 (0.15)		1.86 (0.11)	1.40	63	.17	
P_Lex		1.70 (0.61)		1.16 (0.41)	4.16	51.54	.00	0.89

4.2.5.3 『CEFR-J Wordlist Version 1.0』の語彙レベルからの分析

表 4-9 では、『CEFR-J Wordlist Version 1.0』による語彙レベル別使用率および石川 (2005) で使用されている手法に基づき、中低頻度語彙使用率 (A2-B2 占有率)、低頻度語彙使用率 (B1-B2 占有率) およびレベルインデックスを算出し、習熟度による中低頻度語彙・低頻度語彙使用率・レベルインデックスの各平均値を示した。

表 4-9 習熟度による語彙レベル別使用率・中低頻度語彙・低頻度語彙使用率・レベルインデックス

	A1	A2	B1	B2	A2+B1+B2	B1+B2	Lev Index
C1 レベル	70.81	19.25	6.93	3.01	29.19	9.94	142.15
B2 レベル	74.30	18.74	5.04	1.92	25.70	6.96	134.58
全 体	72.63	18.98	5.94	2.44	27.37	8.38	138.19

表 4-10 では、習熟度（CEFR-J B2 レベルと CEFR-J C1 レベル）による中低頻度語彙・低頻度語彙使用率・レベルインデックスの各平均値の相違を示した。「4.2.5.2 英語の習熟度による発表語彙の豊かさの指標の比較」では、その数値が高いほど低頻度語彙の使用率が高いとされる P_Lex において、習熟度により有意差が認められた。このことは、中低頻度語彙使用率（A2-B2 占有率）、低頻度語彙使用率（B1-B2 占有率）およびレベルインデックスにおいて、習熟度により有意差のある可能性を示唆している。そこで、英語の習熟度の違いにより中低頻度語彙および低頻度語彙使用率、レベルインデックスの各指標に差があるかどうか調べるために、指標ごとに、その平均値の差を対応なしの t 検定により分析した。 t 検定にあたり、いずれの指標に関しても正規性において問題はなかった。また、等分散性を確認するために Levene の検定を行った結果、いずれの指標においても等分散していることが確認できた。 t 検定の結果、中低頻度語彙使用率（A2-B2 占有率）に関しては、習熟度により有意差がみられた ($t(63) = 3.17, p < .01$)。その効果量 Δ は 0.81 となり、効果が大きいこともわかった。低頻度語彙使用率（B1-B2 占有率）に関しては、習熟度により有意差がみられた ($t(63) = 3.60, p < .01$)。その効果量 Δ は 0.80 となり、効果が大きいこともわかった。レベルインデックスに関しては、習熟度により有意差がみられた ($t(63) = 3.54, p < .01$)。その効果量 Δ は 0.86 となり、効果が大きいこともわかった。

表 4-10 習熟度による中低頻度語彙・低頻度語彙使用率・レベルインデックスの各平均値の相違

語彙レベル	C1 レベル	B2 レベル	t 値	2つの母平均の差の検定		
	$N=31$ $M(SD)$	$N=34$ $M(SD)$		自由度	有意確率 (両側)	効果量 Δ
A2+B1+B2	29.19(4.31)	25.70(4.54)	3.13	63	.00	0.81
B1+B2	9.94(3.73)	6.96(2.93)	3.60	63	.00	0.80
Lev Index	142.15(8.87)	134.58(8.34)	3.54	63	.00	0.86

4.2.6 考察

研究課題(1)の「上級英語学習者が記述した自由英作文の全体的評価とそこで使用された語彙の豊かさ(多様性と広範さ)には相関関係があるのか」の解答は、4.2.5の結果より、自由英作文の全体的評価と語彙の多様性とはほとんど相関関係がなく、広範さとは正の中程度の相関関係があるとなった。このことは、語彙の多様性が増しても、英文の質に対する評価者の印象度はあまり向上せず、語彙の広範さが増し、低位頻度語彙の使用率が上がるとその印象度が向上する可能性を示唆している。

海外の先行研究の Engber (1995) では、語彙の多様性 (lexical variation with error と lexical variation without error) と自由英作文の全体的評価との間には正の中程度の相関関係があるという指摘をしており、また、Daller & Phelan (2007) では、語彙の多様性 (GI) と広範さ (P_Lex) の両者と自由英作文の全体的評価との間には正の中程度の相関関係があると指摘しているが、本研究の結果は、これらの先行研究を支持するものではなかった。

CEFR-J A2 レベルの学習者を対象とした 3 章 2 節では、自由英作文の全体的評価と語彙の多様性との間には中程度の相関関係があり、語彙の広範さとは弱い正の相関関係がみられたため、本研究の結果とは異なることになる。3 章 2 節と本研究の相違は、学習者の英語力の相違による語彙力の違いから生じていると考えられる。3 章 2 節では、学習者の英語力が CEFR-J A2 レベルのため、メンタルレキシコンが十分発達しておらず、『CEFR-J Wordlist Version 1.0』の A1 レベルの高頻度語彙が 90%以上を占めており、中・低頻度語彙がほとんど使われていない(9%程度)。このため、90%以上を占める語彙の中で多様性が増すことが評価者の英文の質に対する印象度を高め、低頻度語彙(3%程度)はあまりに使われる量が少ないため印象度の向上につながらなかったと考えられる。

これに対して、本研究では、表 4-9 に示したように、A1 レベルの高頻度語彙が 72%程度で、中・低頻度語彙は 28%程度である。この中・低頻度語彙がより多く使われると評価者の印象度がよくなると考えられる。つまり、異なり語の増加による多様性の向上よりも、中・低頻度語彙の使用による広範さが増すことにより、評価者に英文の質が向上する印象を与えることになると言えよう。

研究課題(2)の「上級英語学習者が記述した自由英作文で使用された語彙の多様性と広範さには相関関係があるのか」の解答は、4.2.5の結果より、中程度の正の相関関係があるとなった。このことは、語彙の多様性と広範さには重なりはあるものの、語彙の豊かさの異なる部分を担っていることの証左とも言えよう。これは、CEFR-J B1 レベルおよび B2 レベルの学習者を対象とした小島(2013)の結果とも一致する。CEFR-J A2 レベルの学習者を対象とした 3 章 2 節では、語彙の多様性と広範さについては正の低い相関関係が認められるのみであった。これらの結果から、CEFR-J B1 レベル以上の英語力の学習者によって記述された自由英作文で使われる語彙の多様性と広範さの相関関係は CEFR-J A2 レベルの場

合よりも向上し、両者の重なりは増すことが示唆される。

研究課題(3)の「英語の習熟度(2群:CEFR-J B2 レベルと CEFR-J C1 レベル)が異なると、語彙の豊かさには差異があるのか」の解答は、4.2.5 の結果より、語彙の多様性に関しては習熟度の差により有意差はみられず、広範さに関しては習熟度の差により有意差がみられるとなった。従って、語彙の多様性よりも語彙の広範さの方が、CEFR-J B2 レベルと CEFR-J C1 レベルの習熟度の異なる学習者で、明確に異なることが明らかとなった。

小島(2013)では、英語の習熟度(2群:CEFR-J B1 レベル下位者と CEFR-J B1 上位者および CEFR-J B2 レベル)の相違により、本研究と同様に、語彙の多様性よりも語彙の広範さの方が明確に異なることを明らかにしている。

また、CEFR-J A2 レベルを対象とした3章2節では、A2 レベルの上位者(TOEIC 401点~550点:英検2級相当)、中位者(TOEIC 351点~400点:英検準2級相当)、下位者(TOEIC 225点~350点:英検3級相当)の3群において、同様の調査を行った結果、多様性に関して、上位と中・下位者の習熟度の相違により、有意差がみられ、一方、広範さに関しては、有意差がみられなかった。従って、CEFR-J A2 レベル内においては、TOEIC 400点を閾値とする習熟度の異なる学習者で、語彙の多様性の方が語彙の広範さよりも明確に異なることになる。

これらの結果から、CEFR-J A2 レベルまでは、習熟度が異なると、語彙の多様性において差異が見られるが、CEFR-J B1 を超えると、習熟度が異なると語彙の広範さにおいて差異がみられることになる。このことは、CEFR-J A2 レベル内の英語力では、上位者といえども、中・低頻度語彙を自由英作文であまり使用できないため、広範さにおいては習熟度による差異はみられず、高頻度語彙の中での異なり語をより使用することによって、多様性に差異が生じると考えられる。CEFR-J B1 レベルを超えると、習熟度が高まるにつれて、中・低頻度語彙を自由英作文でより使用できるようになり、広範さにおいて習熟度の差異がみられるようになると考えられる。また、多様性に関しては、高頻度、中低頻度語彙のいずれにおいても、習熟度が向上しても、異なり語の使用率は必ずしも向上しないため、習熟度により差異がみられないと考えられる。

研究課題(4)の「上級英語学習者が記述した自由英作文では、CEFR-J の語彙リスト『CEFR-J Wordlist Version 1.0』のどのレベルの語彙を使用しているのか」の解答は、4.2.5 の結果から、以下ようになった。

- (1) A1 レベル(高頻度語彙)、すなわち、小学校~中学校2年程度の語彙の使用率が70%を超えていること。
- (2) A2 レベル(中頻度語彙)の中学校3年~高校1/2年程度の語彙の使用率が20%程度であること。
- (3) 高校3年以上のB1 および B2 レベルの語彙(低頻度語彙)は8%程度であること。

同様の調査を実施したCEFR-J A2 レベルの学習者を対象にした3章2節と比較してみる。A1 レベルの高頻度語彙に関して、CEFR-J A2 レベルの学習者では90%を超える使用率で

あった。本研究ではそれが、70%程度となり、20%程度、減少している。当初、本研究での学習者の CEFR-J B2 レベル以上の高い英語力を考慮すると、A1 レベルの高頻度語彙の使用率は、もう少し低い（例えば、60%程度）ことを予測していた。70%程度が A1 レベルの高頻度語彙であるということは、たとえ英語力は高くとも、高頻度語彙を一定量使用しないと、自由英作文というテキストを構築することができないことを表していると考えられる。資料 1-2 に CEFR-J B2 レベルと C1 レベルの自由英作文を提示したが、ここでも A1 レベルの高頻度語彙が相当量使用されていることは明らかである。

A2 レベル以上の中低頻度語彙に関して、CEFR-J A2 レベルの学習者では 9%程度の使用率であったのに対して、本研究では、中低頻度語彙の使用率は 28%程度となっている。CEFR-J A2 レベルと本研究の CEFR-J B2 レベル以上の英語力の相違が、この中低頻度語彙の使用率に反映していると言える。

研究課題(5)の『CEFR-J Wordlist Version 1.0』では英語の習熟度(CEFR-J B2 レベルおよび CEFR-J C1 レベルにより、語彙レベルの使われ方には差があるのか)の解答は、4.2.5の結果から、CEFR-J B2 レベルと CEFR-J C1 レベルの習熟度の違いにより中低頻度語彙使用率、低頻度語彙使用率、レベルインデックスのいずれも有意差があるとなった。英語の習熟度が高くなると、自由英作文においてより低頻度語彙の使用ができることを表している。このことは、表 4-8 における低頻度語彙の使用率が高いと数値が高くなる P_Lex において、CEFR-J C1 レベルと CEFR-J B2 レベルの P_Lex の平均値に有意差のあることと一致する。

4.2.7 結論

4章2節では、CEFR-J B2およびC1レベルの英語力の学習者によって記述された自由英作文を語彙の豊かさを多様性と広範さという観点から分析した。従来、CEFR-J B2およびC1レベルの日本人学習者によって記述された自由英作文の語彙の分析をほとんど行われてこなかった。そのため、CEFR-J B2およびC1レベルのような上級英語学習者によって記述された自由英作文の語彙の豊かさは判明していなかったが、以下の事項が明らかとなった。

- (1) 語彙の多様性と自由英作文の全体的評価とはほとんど相関関係がみられないため、語彙の多様性は自由英作文の言語的特徴とならないことが示唆される。
- (2) 語彙の広範さと自由英作文の全体的評価とは中程度の相関関係がみられ、このことは語彙の広範さが自由英作文の言語的特徴となることを示唆している。
- (3) 日本人上級英語学習者の記述する語彙の多様性と広範さは、語彙の豊かさの異なる側面ではあるが、両者には、中程度の相関関係が存在する。
- (4) 日本人上級英語学習者の記述する自由英作文においては、習熟度が高まるにつれて、多様性には差異が生じないが、広範さにおいては差異が生じる。

(5) 日本人上級英語学習者の記述する自由英作文において使用される語彙は、語彙リスト『CEFR-J Wordlist Version 1.0』の高頻度語彙が70%程度を占め、30%程度が中・低頻度語彙である。

(6) 語彙リスト『CEFR-J Wordlist Version 1.0』においても習熟度が高まるにつれて、中・低頻度語彙の使用率に差異が生じる。

これら分析の過程で、語彙の多様性に関する指標であるGIと語彙の広範さに関する指標であるSはCEFR-J B2およびC1レベルの自由英作文の語彙の調査には適切でないことも判明した。従って、2章5節3項の「CEFR-J B2およびC1レベルの自由英作文の評価の観点となる特定要因の評価指標候補」の語彙の視点の説明変数候補から、GIとSを除外することとした。

4.3 重回帰分析により抽出した CEFR-J B2 および C1 レベルの自由英作文の評価の観点

4.3.1 はじめに

4章3節では、本研究の主目的たる CEFR-J B2 および C1 レベルの自由英作文のルーブリックにおける評価の観点と評価基準（到達度合い）の基礎データを決定する。評価の観点は、CEFR-J B2 および C1 レベルの自由英作文の言語的特徴を重回帰分析によって抽出された説明変数である特定要因の評価指標を使用する。評価基準（到達度合い）は、説明変数となった特定要因の評価指標と自由英作文の全体的評価を基に統計的に検証した定量的数値により、3段階に区分することとする。

4.3.2 研究の目的

4章3節の目的は、CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文の言語的特徴と自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値を明らかにすることである。この目的のために以下の研究課題を設定する。

(1) CEFR-J B2 および C1 レベルの学習者が記述する自由英作文の言語的特徴として有効な説明変数は何か。

(2) 自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値は何か。

(1)によって抽出された言語的特徴として有効な説明変数は CEFR-J B2 および C1 レベルの自由英作文のルーブリックにおける評価の観点となり、(2)の定量的な数値は評価基準（到達度合い）を決定する基礎データとなる。

本研究では、1つの課題の自由英作文で言語的特徴である説明変数を分析していることに疑問を抱かれることを考慮し、統計的手法を駆使することを徹底し、出来る限り一般化することに注意を払った。

4.3.3 方法

4.3.3.1 CEFR-J B2 および C1 レベルの調査参加者

調査参加者は4章1節と同一である（表4-2参照）。

4.3.3.2 ライティングテスト

ライティングテスト4章1節と同一である（表4-3参照）。

4.3.3.3 CEFR-J B2 および C1 レベル自由英作文の評価の観点となる特定要因の評価指標候補

CEFR-J B2 および C1 レベルの自由英作文の言語的特徴に関する特定要因の評価指標の選択に関しては、先行研究を参考にしつつ、構成、内容、言語使用、Fluency、文、語彙、リーダビリティの7つの視点から15変数とした（2章5節3項参照）。しかしながら、4章1

節の結果から、CEFR-J B2 および C1 レベルの自由英作文を評価するリーダビリティの指標は CLI と ARI を妥当とし、FKG を除くこととした。また、4 章 2 節の結果から、語彙の多様性に関する指標である GI と語彙の広範さに関する指標である S は CEFR-J B2 および C1 レベル自由英作文の語彙の調査には適切でないことが判明したため、除外することとした。従って、CEFR-J B2 および C1 レベル自由英作文の評価の観点となる特定要因の評価指標候補は構成、内容、言語使用、Fluency、文、語彙、リーダビリティの 7 つの視点から以下の 12 変数となった。

- (1) 構成の視点から、上級英語学習者によって記述された自由英作文を考慮して、その「構成」を表す特定要因の評価指標を独自に作成した（詳細は 2 章 5 節 3 項参照）。
- (2) 内容の視点から、上級英語学習者によって記述された自由英作文を考慮して、その「内容」を表す特定要因の評価指標を独自に作成した（詳細は 2 章 5 節 3 項参照）。
- (3) 言語使用の視点から、上級英語学習者によって記述された自由英作文を考慮して、その「言語使用」を表す特定要因の評価指標を独自に作成した（詳細は 2 章 5 節 3 項参照）。
- (4) Fluency の視点からは総語数および文数を説明変数候補とした。
- (5) 文の視点からは 1 文の平均語数を説明変数候補とした。
- (6) 語彙の視点からは lexical variation（語彙の多様性）を表す D と MTL D、lexical sophistication（語彙の広範さ）を表す P_Lex、平均語長を説明変数候補とした。
- (7) リーダビリティの視点から ARI、CLI を説明変数候補とした。

表 4-11 に CEFR-J B2 および C1 レベル自由英作文の評価の観点となる特定要因の評価指標候補の記述統計を示す。

表 4-11 CEFR-J B2 および C1 レベル自由英作文の評価の観点となる特定要因の評価指標候補の記述統計 ($N = 65$)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
構成	3.63	0.70	2.00	4.00	2.00	-1.63	1.11
内容	4.11	1.20	0.00	6.00	6.00	-0.87	0.90
言語使用	4.28	1.07	1.00	5.00	4.00	-1.46	1.42
総語数	229.32	42.20	172.00	358.00	186.00	0.70	0.13
文数	12.35	2.61	7.00	19.00	12.00	0.33	-0.11
1 文平均語数	18.98	3.49	13.32	27.42	14.10	0.65	0.10
D	57.77	17.05	29.51	99.99	70.48	0.74	0.08
MTLD	78.71	25.76	39.59	204.55	164.96	2.04	7.84
MTLD#	1.88	0.13	1.60	2.31	0.71	0.44	1.24
P_Lex	1.41	0.58	0.38	2.73	2.35	0.46	-0.36

平均語長	4.96	0.26	4.36	5.84	1.48	0.41	1.09
ARI	11.18	2.38	6.50	17.50	11.00	0.42	0.11
CLI	11.45	1.65	7.90	17.10	9.20	0.41	1.27

表 4-11 の変数のうち、MTLD に関しては正規性に問題があったため、数値変換（MTLD#）を行った。

4.3.3.3 手続き

(1) 自由英作文のテキストデータの収集

調査協力者 65 名に記述された自由英作文のテキストを MS-Word により、デジタルデータ化した。

(2) 全体的評価

自由英作文を ‘Independent WRITING Rubrics’ を基に HE により採点した。

(3) 言語的特徴を表す要因の抽出とその算出

自由英作文の HE に対する説明変数候補を構成、内容、言語使用、Fluency、文、語彙、リーダビリティの 7 つの視点から 12 変数を選択し、その数値を算出した。

(4) (3) の 12 変数を説明変数候補とし、(2) の HE を従属変数として、重回帰分析を行う。

(5) (4) で得られた各説明変数に関して、HE の平均値 4.03(表 4-3 参照)以上を上位者、平均値 4.03 に満たない下位者の 2 群に分け、その平均値に有意差があるかどうか調査する。

(6) (5) の平均値に有意差があれば、各説明変数に関して以下のように仮に定義する。

1. 上位者の平均値以上の数値をその説明変数に関する仮の上位者の定量的な数値とする。
2. 上位者と下位者の平均値の中間の数値をその説明変数に関する仮の中位者の定量的な数値とする。
3. 下位者の平均値以下の数値をその説明変数に関する仮の下位者の定量的な数値とする。

(7) (6) の仮の定量的な数値を以下のように検証する。

1. 2 条件（HE の上・下位者）×3 値（各説明変数の上・中・下位者）のクロス集計表を作成する。
2. 上記 1 のクロス集計表を基に、 χ^2 検定および残差分析を行う。

4.3.4 結果

4.3.4.1 重回帰分析の結果

自由英作文の全体的評価（HE）を従属変数とし、12 変数を説明変数候補とする重回帰分析を行った。重回帰分析はステップワイズ法を採用した。表 4-12 はその結果を示し、Model

1 が従属変数である HE の予測に役立つことも判明した ($F = 55.34$, $p < .001$)。HE の要因として抽出された説明変数は内容、言語使用、構成の 3 変数である。これらの 3 説明変数による説明率は約 73% ($R^2 = 0.73$) である。

説明変数間の関係に関しては、内容と言語使用、内容と構成、言語使用と構成のピアソンの相関係数がそれぞれ、0.34 ($p < .01$ 両側検定)、0.15、-0.09 となっており、内容と言語使用とは低い正の相関関係があるが、他の変数間ではほとんど相関関係がみられない係数となっている (表 4-13 参照)。内容と言語使用の間には低い正の相関関係はあるものの、 $R^2 = 0.73$ に対し、調整済み $R^2 = 0.71$ であり、ほとんど数値が下がっていない。従って、これらの説明変数間は互いに独立しているとみなしても差し支えないと考えられる。

多重共線性を判断する指標として VIF ($VIF = 1 / \text{許容度}$) がある。「VIF が 10 未満である」、「VIF の平均値が 1 から大きく離れていない」という条件が満たされれば、多重共線性ではないと判断する (竹内・水本, 2012, p. 138)。構成の許容度 (0.95)・VIF (1.05)、内容の許容度 (0.85)・VIF (1.18)、言語使用の許容度 (0.86)・VIF (1.16)、となっており、多重共線性になっていないと判断される。標準化係数 β 値から判断すると、構成、内容、言語使用、はそれぞれ 0.20:0.49:0.51 の比率で HE に貢献することが判明した。

表 4-12 重回帰分析の結果

Model 1	非標準化係数		標準化係数			許容度	VIF
	R^2	B	標準誤差	β	t 値		
定数項		.91	.30		2.98	.00	
構成	.05	.19	.06	.20	2.95	.00	.95
内容	.49	.27	.04	.49	6.86	.00	.85
言語使用	.44	.31	.04	.51	7.16	.00	.86

$R = .86$ $R^2 = .73$ 調整済み $R^2 = .72$

表 4-13 相関行列表

	(1)	(2)	(3)	(4)
(1) HE	—	.23	.70**	.66**
(2) 構成		—	.15	-.09
(3) 内容			—	.34**
(4) 言語使用				—

** : $p < .01$

この重回帰分析の結果は、CEFR-J B2 および C1 レベルの自由英作文のルーブリックの作成に有用な基礎データになると考えられる。この Model 1 は決定係数 R^2 も 0.73 となって

おり、抽出された3つの説明変数によって全体的評価の73%が説明されるというのは優れたモデルであると言えよう。しかしながら、抽出された3つの説明変数は評価者によって作成される^{注31)}ものであり、例えば、総語数のようなより客観性の高い変数が抽出される可能性はないかと考えたのである。そのような説明変数を探るとともに、2章5節1項に記述したように、自由英作文の言語的特徴に関する先行研究から少なくとも決定係数 R^2 が最低限0.50を超えることは条件としてみた。

結果として、以下のようなModel 2となった。表4-14はその結果を示し、Model 2が従属変数であるHEの予測に役立つことも判明した($F = 36.44$, $p < .001$)。HEの要因として抽出された説明変数は言語使用とCLIの2変数である。これらの2説明変数による説明率は約54% ($R^2 = 0.54$)である。言語使用は語彙や文法の誤りに関して、エラーの数で5段階の得点を付与する(2章5節3項参照)ことによって算出する特定要因の評価指標であり、ローカルエラーまでを誤りとしたので、客観性が高い指標である。また、CLIは計算用のプログラムによって算出されるため、客観性の高い指標である。

説明変数間の関係に関しては、言語使用とCLIのピアソンの相関係数が、0.24となっており、低い正の相関関係(但し、 $p > .05$)がある(表4-15参照)。しかしながら、 $R^2 = 0.54$ に対し、調整済み $R^2 = 0.53$ であり、ほとんど数値が下がっていない。従って、これらの説明変数間は互いに独立しているとみなしても差し支えないと考えられる。

多重共線性を判断する指標としてVIF ($VIF = 1 / \text{許容度}$)がある。「VIFが10未満である」、「VIFの平均値が1から大きく離れていない」という条件が満たされれば、多重共線性ではないと判断する(竹内・水本, 2012, p. 138)。言語使用の許容度(0.95)・VIF(1.06)、CLIの許容度(0.95)・VIF(1.06)となっており、多重共線性になっていないと判断される。標準化係数 β 値から判断すると、言語使用とCLIは0.59:0.33の比率でHEに貢献することが判明した。

表 4-14 重回帰分析の結果

Model 2	非標準化係数			標準化係数			許容度	VIF
	R^2	B	標準誤差	β	t 値	p		
定数項		1.02	.42		2.44	.02		
言語使用	.44	.36	.54	.59	6.62	.00	.95	1.06
CLI	.22	.13	.04	.33	3.69	.00	.95	1.06

$R = .74$ $R^2 = .54$ 調整済み $R^2 = .53$

^{注31)} CEFR-J B2 および C1 レベルのような上級者によって記述される自由英作文には評価者によって採点される部分は必要と認めたので、構成、内容、言語使用という評価指標を作成した。

表 4-15 相関行列表

	(1)	(2)	(3)
(1) HE	—	.66**	.47**
(2) 言語使用		—	.24
(3) CLI			—

**： $p < .01$

4.3.4.2 説明変数の定量的な数値

4.3.3.3 の手続き (5)～(7)について Model 1 と Model 2 のモデルの結果を以下に示す。Model 1 の説明変数である内容、言語使用、構成について、HE の上位者 29 名と下位者 36 名に関して、独立 2 群の平均値の差の検定を行う。この場合、それぞれの標本の正規分布と等分散性という 2 つの条件を満たさなければならない。抽出された 3 説明変数に関して、正規分布は満たしていたが、Levene 検定により、いずれの変数も等分散性でないことが判明した（構成： $F = 49.24$, $p = .00$ ；内容： $F = 8.98$, $p = .00$ ；言語形式： $F = 50.93$, $p = .00$ ）。従って、3 説明変数に関しては Welch の t 検定を行うこととした（表 4-16 参照）。

構成に関しては、Welch の t 検定の結果、HE の上・下位者の平均値の差は有意であった（両側検定： $t(46) = 3.17$, $p = .00$ ）。また、 t 検定の効果量に関しては、Hedges の g を使用することとし、効果量の目安は「0.20 (小)、0.50 (中)、0.80 (大)」(平井, 2012, p. 48) に従った。Hedges の効果量を算出した結果、 $g = 0.42$ となり、小から中の効果があることがわかった。

内容に関しては、Welch の t 検定の結果、HE の上・下位者の平均値の差は有意であった（両側検定： $t(57) = 6.23$, $p = .00$ ）。Hedges の効果量を算出した結果、 $g = 0.64$ となり、中から大の効果があることがわかった。

言語使用に関しては、Welch の t 検定の結果、HE の上・下位者の平均値の差は有意であった（両側検定： $t(37) = 6.32$, $p = .00$ ）。Hedges の効果量を算出した結果、 $g = 0.72$ となり、中から大の効果があることがわかった。

表 4-16 内容・言語使用・構成に関する HE の上・下位者の平均値

言語的特徴を表 す変数	HE 上位者		HE 下位者		2 つの母平均の差の検定			
	$N = 29$	$M(SD)$	$N = 36$	$M(SD)$	t 値	自由度	有意確率 (両側)	効果量 g
構成	3.90	(.31)	3.42	(.84)	3.17	46.16	.00	.42
内容	4.90	(.66)	3.47	(1.16)	6.23	57.26	.00	.64
言語使用	4.97	(.19)	3.72	(1.16)	6.32	37.21	.00	.72

このように各変数に関して、HE の上・下位者の平均値には有意差が認められたため、仮

の定義として、変数の上位者の平均値以上（構成：3.90以上；内容：4.90以上；言語使用：4.97以上）、上位者と下位者の平均値の間（構成：3.89～3.43；内容：4.89～3.48；言語使用：4.96～3.73）、下位者の平均値以下（構成：3.42以下；内容：3.47以下；言語使用：3.72以下）に分けてみた。さらに、HEの上・下位者の人数と各変数の上・中・下位者の人数の2条件×3値でクロス集計し、 χ^2 検定を行い、有意差があるかどうか検証した。

表4-17はHEの上・下位者と構成の上・中・下位者のクロス集計表である。 χ^2 検定の効果量については、CramerのVに関する一般的基準「0.1（効果量小）、0.3（効果量中）、0.5（効果量大）」に従う（竹内・水本，2012，p.353）。 χ^2 検定の結果、人数の偏りは有意傾向（ $\chi^2(2) = 5.89, p = .05$ ）となったが、CramerのVが0.30となり、効果量は中程度となったことから、十分に意味のある差であると考えられる。従って、残差分析の結果、構成が3.42以下だと、HEの下位者は増え、構成が3.90以上だと、HEの上位者が増えることと判断し、構成に関して、仮の定義は検証された。

表4-18はHEの上・下位者と内容の上・中・下位者のクロス集計表である。 χ^2 検定の結果、人数の偏りは有意である（ $\chi^2(2) = 26.44, p = .00$ ）。また、効果量としてCramerのVが0.64となったため、効果量大と判断される。残差分析の結果、内容が3.47以下だと、HEの下位者は増え、内容が4.90以上だと、HEの上位者が増えることも判明した。従って、内容に関して、仮の定義は検証された。

表4-19はHEの上・下位者と言語使用の上・中・下位者のクロス集計表である。 χ^2 検定の結果、人数の偏りは有意である（ $\chi^2(2) = 29.33, p = .00$ ）。また、効果量としてCramerのVが0.67となったため、効果量大と判断される。残差分析の結果、言語使用が3.72以下だと、HEの下位者は増え、言語使用が4.97以上だと、HEの上位者が増えることも判明した。従って、言語使用に関して、仮の定義は検証された。

表4-17 HEの上・下位者と構成の上・中・下位者のクロス集計表

		構 成			合計
		上位者 3.90 以上	中位者 3.89～3.43	下位者 3.42 以下	
HE	度数	26	0	3	29
上位者	調整済み残差	2.4	-.9	-2.2	
HE	度数	23	1	12	36
下位者	調整済み残差	-2.4	.9	2.2	
合計	度数	49	1	15	65

表 4-18 HE の上・下位者と内容の上・中・下位者のクロス集計表

		内 容			合計
		上位者	中位者	下位者	
		4.90 以上	4.89～3.48	3.47 以下	
HE	度数	19	9	1	29
上位者	調整済み残差	4.8	-1.5	-3.6	
HE	度数	3	18	15	36
下位者	調整済み残差	-4.8	1.5	3.6	
合計	度数	22	27	16	65

表 4-19 HE の上・下位者と言語使用の上・中・下位者のクロス集計表

		言語使用			合計
		上位者	中位者	下位者	
		4.97 以上	4.96～3.73	3.72 以下	
HE	度数	28	1	0	29
上位者	調整済み残差	5.4	-2.8	-3.8	
HE	度数	11	11	14	36
下位者	調整済み残差	-5.4	2.8	3.8	
合計	度数	39	12	14	65

次に、Model 2 の説明変数である言語使用、CLI について、HE の上位者 29 名と下位者 36 名に関して、独立 2 群の平均値の差の検定を行う。言語使用に関しては、Model 1 でも説明変数となっており、表 4-18 で定量的な数値を提示してある。CLI は正規分布と等分散性の 2 つの条件を満たしているため、 t 検定を行うこととした（表 4-20 参照）。

CLI に関しては、の t 検定の結果、HE の上・下位者の平均値の差は有意であった（両側検定： $t(63) = 3.81$, $p = .00$ ）。Hedges の効果量を算出した結果、 $g = 0.43$ となり、小から中の効果があることがわかった。

表 4-20 CLI に関する HE の上・下位者の平均値

言語的特徴を表 す変数	HE 上位者		HE 下位者		2 つの母平均の差の検定			
	$N = 29$	$M(SD)$	$N = 36$	$M(SD)$	t 値	自由度	有意確率（両側）	効果量 g
CLI		12.24 (1.58)		10.82 (1.42)	3.81	63.00	.00	.43

このように CLI に関して、HE の上・下位者の平均値には有意差が認められたため、仮の定義として、変数の上位者の平均値である 12.24 以上、上位者と下位者の平均値の間で

ある 12.23~10.83、下位者の平均値である 10.82 以下に分けてみた。さらに、HE の上・下位者の人数と各変数の上・中・下位者の人数の 2 条件×3 値でクロス集計し、 χ^2 検定を行い、有意差があるかどうか検証した。

表 4-21 は HE の上・下位者と CLI の上・中・下位者のクロス集計表である。 χ^2 検定の結果、人数の偏りは有意である ($\chi^2(2) = 9.89, p = .01$)。また、効果量として Cramer の V が 0.39 となったため、効果量中と判断される。残差分析の結果、CLI が 10.82 以下だと、HE の下位者は増え、CLI が 12.24 以上だと、HE の上位者が増えることも判明した。従って、CLI に関して、仮の定義は検証された。

表 4-21 HE の上・下位者と CLI の上・中・下位者のクロス集計表

		CLI			合計
		上位者 12.24 以上	中位者 12.23~10.83	下位者 10.82 以下	
HE	度数	13	10	6	29
上位者	調整済み残差	2.8	.1	-2.6	
HE	度数	5	12	19	36
下位者	調整済み残差	-2.8	-.1	2.6	
合計	度数	18	22	25	65

4.3.5 考察

研究課題 (1) 「CEFR-J B2 および C1 レベルの学習者が記述する自由英作文の言語的特徴として有効な説明変数は何か」については、重回帰分析の結果、2 つのモデルが得られたので、2 種類の解答が得られた。Model1 での説明変数は内容、言語使用、構成である。これらの 3 変数は自由英作文のルーブリックの評価の観点になる可能性がある。標準化係数 β 値は構成が 0.20、内容が 0.49、言語使用が 0.51 となり、内容と言語使用の自由英作文に対する HE に対して、ほぼ同じ比率で、また、構成はそれらの両変数の約 40% の割合で HE に貢献していることが判明した。従って、 β 値の比率はほぼ、構成：内容：言語使用：=2:5:5 となる。

Model1 の説明率は 73% 程度となり、先行研究の Ferris (1994) の 50% 程度、杉浦 (2008) の 44% 程度、水本 (2008) の 60% 程度、4 章 3 節の CEFR-J A2 レベルの自由英作文の言語的特徴の 73% 程度と比較しても、有効なモデルであることを示している。

これらの抽出された 3 説明変数をもとに、日本人の上級英語学習者に自由英作文を指導する際の評価に関する示唆を以下にまとめる。

- (1) 構成として、序論において与えられたトピックを踏まえ、本論につながる構成をし、また、結論において本論と首尾一貫したつながりを有している自由英作文の評価は高

くなる可能性がある。

(2) 内容としての本論の主題部分と支持部分で評価の高い自由英作文は評価が高くなる可能性がある。

(3) 言語使用としての語彙や文法の誤りの少ない自由英作文は評価が高くなる可能性がある。

Model2 での説明変数は言語使用と CLI である。これらの 2 変数は自由英作文のルーブリックの評価の観点になる可能性がある。標準化係数 β 値は言語使用が 0.59、CLI が 0.33 となり、ほぼ、2:1 (正確には 1.8 :1) の比率で HE に貢献していることが判明した。

Model2 の説明率は 54% 程度となり、Model1 の説明率 73% には及ばぬものの、言語使用と CLI というより客観性の高い変数が得られるという長所がある。

これらの抽出された 2 説明変数をもとに、日本人の上級英語学習者に自由英作文を指導する際の評価に関する示唆を以下にまとめる。

(1) 言語使用としての語彙や文法の誤りの少ない自由英作文は評価が高くなる可能性がある。

(2) CLI の数値の高いリーダビリティの優れた作文は高評価を受ける可能性が高くなる。

研究課題 (2) 「自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値は何か」については、Model1 から得られた解答として構成「上位者:3.90 以上; 中位者: 3.89~3.43 ;下位者: 3.42 以下」、内容「上位者:4.90 以上;中位者: 4.89~3.48 ; 下位者: 3.47 以下」、言語使用「上位者:4.97 以上;中位者: 4.96~3.73;下位者: 3.72 以下」となる。このような数値化は自由英作文のルーブリックの 3 段階の評価基準 (到達度合い) を決定する基礎データとなる。表 4-22 に CEFR-J B2 および C1 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値を示した。

表 4-22 Model1 から得られた CEFR-J B2 および C1 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値

	上位者	中位者	下位者
構成	3.90 以上	3.89~3.43	3.42 以下
内容	4.90 以上	4.89~3.48	3.47 以下
言語使用	4.97 以上	4.96~3.73	3.72 以下

Model2 から得られた解答として言語使用「上位者:4.97 以上;中位者: 4.96~3.73;下位者: 3.72 以下」、CLI「上位者: 12.24 以上;中位者: 12.23~10.83 ;下位者: 10.82 以下」となる。このような数値化は自由英作文のルーブリックの 3 段階の評価基準 (到達度合い) を決定する基礎データとなる。表 4-23 に CEFR-J B2 および C1 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値を示した。

表 4-23 Model2 から得られた CEFR-J B2 および C1 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値

	上位者	中位者	下位者
言語使用	4.97 以上	4.96～3.73	3.72 以下
CLI	12.24 以上	12.23～10.83	10.82 以下

4.3.6 結論

4章3節では、CEFR-J B2 および C1 レベルの自由英作文のルーブリックにおける評価の観点と評価基準（到達度合い）の基礎データを探求した。評価の観点は自由英作文の言語的特徴を重回帰分析による説明変数である特定要因の評価指標により抽出した。評価基準（到達度合い）は説明変数となった特定要因の評価指標に関して、自由英作文の全体的評価を基に、統計的に検証した定量的数値により3段階に区分した。以下に、これらの分析により得られた知見を示す。

- (1) CEFR-J B2 および C1 レベルの自由英作文の言語的特徴は重回帰分析による2つのモデルによる2種類の特定要因の評価指標の組み合わせとなった。Model 1 では、構成、内容、言語使用の3つの特定要因の評価指標である。これらの3変数は自由英作文のルーブリックにおいて評価の観点となる可能性がある。Model 2 では、言語使用、CLIの2つの特定要因の評価指標である。これらの2変数も自由英作文のルーブリックにおいて評価の観点となる可能性がある。
- (2) (1)の抽出された言語的特徴の中で、Model 1の3指標の標準化係数 β 値の比率は、構成：内容：言語使用：=2:5:5となる。
- (3) (1)の抽出された言語的特徴の中で、Model 1の3指標に関して、統計的に検証した結果、3段階の定量的な数値を得ることができた。これによりのCEFR-J B2 および C1 レベル自由英作文のルーブリックにおいて、構成、内容、言語使用に関しては3段階の評価基準（到達度合い）を有する評価の観点とすることが可能であることが明らかとなった。
- (4) (1)の抽出された言語的特徴の中で、Model 2の2指標の標準化係数 β 値の比率は、言語使用：CLI=2:1（正確には1.8:1）となる。
- (5) (1)の抽出された言語的特徴の中で、Model 2の2指標に関して、統計的に検証した結果、3段階の定量的な数値を得ることができた。これによりのCEFR-J B2 および C1 レベル自由英作文のルーブリックにおいて、言語使用、CLIに関しては3段階の評価基準（到達度合い）を有する評価の観点とすることが可能であることが明らかとなった。
- (6) 評価の観点となる構成、内容、言語使用の評価基準（達成度合い）の定量的数値は表4-22のようになった。

(7)評価の観点となる言語使用、CLI の評価基準（達成度合い）の定量的数値は表 4-23 のようになった。

これらの結果を基に、CEFR-J B2 および C1 レベルの自由英作文のルーブリックをデザインする。

4.4 まとめ

4章では、CEFR-J B2 および C1 レベルの自由英作文についてリーダビリティと語彙の豊かさの視点から分析した。その結果を踏まえて、重回帰分析により CEFR-J B2 および C1 レベルの自由英作文の言語的特徴を抽出し、ルーブリックにおける評価の観点を探ることとした。

リーダビリティの視点から分析した結果として、CEFR-J B2 および C1 レベルの自由英作文を読み、理解するにはアメリカの学校制度における 10 年生から 11 年生の教育レベルが必要なこと、また、リーダビリティの視点から分析する際に使用したリーダビリティの指標である ARI と CLI は CEFR-J B2 および C1 レベルの自由英作文の言語的特徴を表わす特定要因の評価指標になりうる可能性があることが判明した。

語彙の豊かさ（多様性と広範さ）の視点から分析した結果として、語彙の多様性よりも広範さの方が CEFR-J B2 および C1 レベルの自由英作文の言語的特徴になりうる可能性が高いことが判明した。多様性と広範さの関係については、両者には重なる部分があり、中程度の相関関係があることが明らかとなった。CEFR-J B2 レベル（TOEIC 785 点～945 点）と CEFR-J C1 レベル（TOEIC 945 点～990 点）という習熟度の相違で、語彙の多様性には差異は生じないが、広範さにおいては差異が生じることも明らかとなった。また、CEFR-J A2 レベルの自由英作文で使用される語彙は『CEFR-J Wordlist Version 1.0』（東京外国語大学投野由紀夫研究室）の A1（小学校～中学校 2 年程度）という高頻度語彙が 70%程度、A2・B1・B2 の中低頻度（A2：中学 3 年～高校 1/2 年程度；B1：高校 3 年～大学受験レベル；B2：大学受験～大学教養レベル）が 30%程度であることも判明した。

CEFR-J B2 および C1 レベルの自由英作文に関して、重回帰分析により言語的特徴を抽出した結果、Model 1 と Model 2 の 2 つのモデルが得られた。Model 1 では、構成、内容、言語使用の 3 つの特定要因の評価指標が得られた。これらの指標は本研究におけるルーブリックの評価の観点となる可能性がある。この 3 指標に関して、ルーブリックの評価基準（到達度合い）の作成ため、統計的に有意になるように 3 段階の定量的な数値の算出を試みた。結果として、これらの 3 指標については統計的に有意な 3 段階の定量的な数値を得られたため、ルーブリックの 3 段階の評価基準（到達度合い）を有する評価の観点として利用できることが判明した。従って、Model 1 による CEFR-J B2 および C1 レベルの自由英作文のルーブリックについては構成、内容、言語使用の 3 つの特定要因の評価指標を評価の観点とすることとなった。

Model 2 では、言語使用、CLI の 2 つの特定要因の評価指標が得られた。これらの指標

は本研究におけるルーブリックの評価の観点となる可能性がある。この 2 指標に関して、ルーブリックの評価基準（到達度合い）の作成ため、統計的に有意になるように 3 段階の定量的な数値の算出を試みた。結果として、これらの 2 指標については統計的に有意な 3 段階の定量的な数値を得られたため、ルーブリックの 3 段階の評価基準（到達度合い）を有する評価の観点として利用できることが判明した。従って、Model 2 による CEFR-J B2 および C1 レベルの自由英作文のルーブリックについては言語使用、CLI の 2 つの特定要因の評価指標を評価の観点とすることとなった。

5. 特定要因の評価指標を用いた評価の観点に基づく自由英作文のルーブリックの提案

5.1 特定要因の評価指標を用いた評価の観点に基づくルーブリックのデザイン方法

5.1.1 ルーブリックのデザインに関する考え方

自由英作文には全体的評価、分析的評価に基づくルーブリック（2章参照）があるが、本研究では、自由英作文の言語的特徴を表わす特定要因の評価指標を評価の観点とするルーブリックをデザインする。ルーブリックの形式は1章2節で提示した形式（表1-7）を採用する。このようなルーブリックをデザインする基本的な考え方は Nishijima, et al (2007) の7項目の質問に対する回答に基づくことにした。これらの質問と回答を表5-1にまとめた。

表 5-1 Nishijima, et al (2007)の質問項目に対する回答

質問番号	回答
(1) 誰がそのルーブリックを使用するのか。	CEFR-J A2 レベルの学習者とその指導者、CEFR-J B2 および C1 レベルの学習者とその指導者
(2) ルーブリックは広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するのに使われるのか、それとも、1つのタスクに対する深い達成度を評価するのに使われるのか。	広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するのに使用する。
(3) ルーブリックは全体的評価とするのか、それとも、分析的評価とするのか。	分析的評価とする。
(4) ルーブリックではどのような得点の付け方をするのか。	質問(5)の評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値を基に設定し、評価基準の得点を重回帰分析の標準化係数 β 値（自由英作文の全体的評価に対する評価の観点の貢献度）の比率により設定する。
(5) 評価の観点には何を使用すべきか。	自由英作文の言語的特徴を重回帰分析により抽出し、それを評価の観点とする。
(6) 評価の観点はどのように定義すべきか。	質問(5)の評価の観点とした言語的特徴の意味合いを考慮して定義する。
(7) その評価の規準（到達の度合い）の水準をどのようにすべきか。	ルーブリックの使用者の使いやすさを考慮し、3水準とする。

表 5-1(1)は、CEFR-J A2 レベルでは一般的な日本人の高校生の英語力であるため、ルーブリックの使用者は高校生とその指導者である高等学校の先生となる。CEFR-J B2 および C1 レベルではルーブリックの使用者は日本人上級英語学習者とその指導者となる。

表 5-1(2)では、汎用的なルーブリックをデザインすることを意図した。

表 5-1(3)は、分析的評価が全体的評価よりも信頼性の高いとされることや学習者に評価項目に関する情報が提供できることを意図した。

表 5-1(4)のような得点の付け方は従来のルーブリックでは採用されていない。

表 5-1(5) のような評価の観点は従来のルーブリックでは採用されていない。

表 5-1(6)では、例えば、評価の観点とした特定要因の評価指標が総語数の場合、自由英作文で記述された量の多さを表わすため、「情報性の豊かさ」という評価の観点とする。

表 5-1(7)について Nishijima, et al (2007)では、3 水準の場合、評価者が中間水準を選びがちなることを懸念しているが、本研究におけるルーブリックでは自由英作文のテキストの客観的かつ信頼性の高い観点（言い換えれば、言語的特徴）の数値に基づくため、評価者による中間水準に集中するという懸念はなくなることになる。

5.1.2 本研究のルーブリックの新規性

従来のルーブリック（2章参照）には、以下の4点で問題があると考えられる。

- (1)自由英作文の言語的特徴を実証した評価の観点が用いられていない。
- (2)評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値となっていない。
- (3)評価の観点に対する評価基準（到達度合い）の得点の設定方法が明確になっていない。
- (4)評価を示す記述（パフォーマンスの特徴を示した記述語）が一般的かつ抽象的な表現が使われて記述されているため、学習者や指導者へのフィードバックが明確でない。

上記(1)～(4)の問題を解決するために、本研究でデザインするルーブリックでは、以下のような新規性がある。

- (1)重回帰分析により抽出した自由英作文の言語的特徴を特定要因の評価指標を評価の観点とする。
- (2)特定要因の評価指標を評価の観点とした分析的評価とする。
- (3)評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値とする。
- (4)評価の観点に対する評価基準（到達度合い）の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値（全体的評価に対する評価の観点の貢献度）の比率により設定する。

(5) 評価を示す記述^{注32)}に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、学習者や指導者に提供することができる。

このような、ルーブリックのデザイン方法を基に以下に CEFR-J A2 レベルの自由英作文のルーブリックと CEFR-J B2 および C1 レベルの自由英作文のルーブリックを提案する。

^{注32)} 1章2節で、ルーブリックを学習者の「パフォーマンスの成功の度合いを示す尺度と、それぞれの尺度に見られるパフォーマンスの特徴を説明する記述語で構成される、評価基準の記述形式」として定義される評価ツールとしているが、この「それぞれの尺度に見られるパフォーマンスの特徴を説明する記述語」に相当する。

5.2 CEFR-J A2 レベルの自由英作文のルーブリックの提案

5.2.1 はじめに

5章2節では、CEFR-J A2 レベルの自由英作文に関する言語的特徴を表わす特定要因の評価指標を評価の観点とする分析的評価に基づくルーブリックを提案する。

ルーブリックに関する先行研究については2章でまとめたが、CEFR-J A2 レベルを対象としたルーブリックはこれまで作成されていない。言い換えれば、自由英作文のような一定量の英文を評価基準の指針が明確に定まっていないことになる。そのため、CEFR-J A2 レベルの一般的な対象となる高校生を指導する高校の英語教師にとっては自由英作文の適正な評価を行うことは非常に苦勞することになる。

CEFR-J A2 レベルというのは自由英作文の始動期と考えられる英語力のため、自由英作文を記述するライティング能力の欠如により、自由英作文自体は稚拙な記述（資料1-1参照）であるため、評価が困難になるのである。既存の自由英作文に関するルーブリックの最も一般的なものの内の1つの‘Independent WRITING Rubrics’では、CEFR-J A2 レベルの大半が1点の範疇にはいり、英文の質に差のある場合でも、評価の峻別ができず、ルーブリックとして活用するには難がある。

このような自由英作文の評価基準の指針が定まっていないことが、高等学校において自由英作文の教育が普及を阻害する一因になっていると考えられる。この状況を克服するためにもCEFR-J A2 レベルのルーブリックをデザインする意義は大きいと考える。

5.2.2 研究の目的

本研究の目的はCEFR-J A2 レベルの自由英作文に関する言語的特徴を表わす特定要因の評価指標を評価の観点とする分析的評価に基づくルーブリックをデザインすることである。そのルーブリックでは、評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値を使用し、評価の観点に対する評価基準（到達度合い）の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値（全体的評価に対する評価の観点の貢献度）の比率により設定する。さらに、このルーブリックでは、評価を示す記述に加え、評価の観点の定量的な数値を掲載する。

5.2.3 研究方法

5.2.3.1 CEFR-J A2 レベルの調査参加者

調査参加者は3章1節と同一である（表3-1参照）。

5.2.3.2 手続き

(1)Nishijima, et al (2007) の7つの質問に対する回答を検討し、ルーブリックの基本設計を行う。

- (2) 3章3節の結果として得られた自由英作文の言語的特徴を表わす特定要因の評価指標を評価の観点とする。
- (3) 評価の観点に対する評価基準（到達度合い）を3章3節で統計的に検証した定量的数値を使用する。
- (4) 評価の観点に対する評価基準（到達度合い）の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値の比率により設定する。
- (5) 評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、一般的で抽象的な評価を示す記述を補完する。

5.2.4 CEFR-J A2 レベルの自由英作文のルーブリックのデザイン

表 5-1 に Nishijima, et al (2007) の質問項目とその回答を提示した。ここでは、Nishijima, et al (2007) の質問に対する回答をより詳細に示し、ルーブリックをデザインする。

「(1) 誰がそのルーブリックを使用するのか」については、本研究の対象は CEFR-J A2 レベルの学習者となり、一般的な日本人の高校生であるため、ルーブリックを使用するのは指導者としての高等学校の先生方を対象とする。また、指導の際に、学習者にルーブリックを提示することも、自由英作文の教育指導に効果を発揮すると考え、高校生の学習者もルーブリックの利用者と考えた。

「(2) ルーブリックは広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するのに使われるのか、それとも、1つのタスクに対する深い達成度を評価するのに使われるのか」については、広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するルーブリックとすることを意図する。この意図を達成するため、CEFR-J A2 レベルの自由英作文を統計的に実証した言語的特徴に関する結果を活用し、汎用化に耐えうるルーブリックのデザインすることとした。

「(3) ルーブリックは全体的評価とするのか、それとも、分析的評価とするのか」については、分析的評価を採用した。これは、分析的評価が全体的評価よりも信頼性の高いとされることや学習者に対して診断的フィードバックを提供するという利点があるためである (Hamp-Lyons, 1991a ; Jacobs et al, 1981)。

「(4) ルーブリックではどのような得点の付け方をするのか」については、4章3節の CEFR-J A2 レベルの言語的特徴に関する実証研究の結果、説明変数となった4変数の内の3変数（総語数、CLI、EFT 平均語数）における重回帰分析の標準化係数（ β 値、表 3-11 参照）の比率と(7)の質問の回答の3水準を活用した。

「(5) 評価の観点には何を使用すべきか」については、3章3節の CEFR-J A2 レベルの英語力の学習者によって記述された自由英作文の言語的特徴を表す特定要因の評価指標である総語数、CLI、EFT 平均語数を使用する。

「(6) 評価の観点はどうのように定義すべきか」については、(5)の回答で提示した総語数、

CLI、EFT 平均語数の意味あいを考慮して定義する。総語数は表 3-11 の標準化係数 β 値において 0.74 であり、CLI や EFT 平均語数に比べて、自由英作文の全体的評価に対する貢献度が圧倒的に高いことが明らかである。このことは、全体的評価は情報量の多さである総語数によって大きく影響を受けていると考えられる。この点を考慮して、ループリックの評価の観点の定義として、総語数については「情報性の豊かさ」とする。

CLI はリーダビリティの指標であり、その算出式は染谷（2009）によれば、次の公式で算出される（3 章 1 節参照）。

$$CLI = (5.89 \times CPW) + (0.3 \times (100 \div WPS)) - 15.80$$

CPW=average number of characters per word WPS=average number of words per sentence
CPW は平均語長、WPS は平均文長を表している。従って、CLI で示されるリーダビリティは高梨・卯城（2000）の「語彙の難易、文長等の文体によるよみやすき」となる。本研究では、CLI の算出公式および高梨・卯城（2000）の定義を考慮して、ループリックの観点の定義として CLI を「語彙や文体の難易度」とする。

EFT 平均語数に関しては、エラーと T-unit の 2 つの視点を考慮することとした。エラーに関しては、「文もしくは発話の理解を難しくするか、もしくは全く理解できなくするような文構造の主要素の使い方の誤り」（Richards et al., 1992）という定義に従い、グローバルエラーを対象とした。また、T-unit に関しては、Hunt（1965）によって考案された概念で、「文の言語的な複雑性を表す尺度で、文が分割される最小単位として定義され、1 つの独立節とそれに付加されるあらゆる従属節とともに成り立っている」（Richards et al., 1992）という定義に従い、その数値が上がれば、書き手は統語的に熟達した作文を書く能力を有することになる（2 章 5 節 2 項参照）。この 2 つの視点を考慮し、ループリックの評価の観点の定義として EFT 平均語数を「エラーが少なく、熟達した英文」とした。

「(7) その観点の水準をどのようにすべきか」については、3 章 3 節で抽出された 3 説明変数（総語数、CLI、EFT 平均語数）に関して上・中・下位者の定量的な数値を算出しており、それをループリックの評価の観点の水準として使用する。

この 7 つの回答に加え、以下に得点算出に関わる部分 (4) と (7) の回答に関して詳述する。(7) で 3 水準を用いることにしたので、各変数の上・中・下位者に対して、仮に 3 点、2 点、1 点を割り当てる。この仮の得点に (4) の回答による考え方を加味する。3 章 3 節の重回帰分析の表 3-11 の総語数、CLI、EFT 平均語数の標準化係数 (β 値) がそれぞれ 0.74、0.19、0.12 となっており、ほぼ 6 : 1.5 : 1 の比率となっている。

重回帰分析における β 値はその説明変数がどの程度の割合で目的変数（自由英作文の全体的評価）に貢献しているかを表す。この場合、総語数が他の説明変数である CLI や EFT 平均語数の平均語数よりも圧倒的で、比率でいえば、上述したように、総語数と CLI は 4 : 1、総語数と EFT 平均語数は 6 : 1、また、CLI と EFT 平均語数は 1.5 : 1 となっている。この各説明変数（ループリックにおける評価の観点）の目的変数に対する貢献度をループリックにおける得点の算出に活用する。

つまり、仮に設定した各変数の上・中・下位者の3点、2点、1点にCLI、EFT平均語数の標準化係数の比率、6：1.5：1を掛け合わせることにする。このようにして、各説明変数の得点を算出したのが表5-2である。従って、この場合の最高点は25.5点で最低点は8.5点となる。

表5-2 各説明変数の上・中・下位者の得点

	上位者	中位者	下位者
総語数	18点	12点	6点
CLI	4.5点	3点	1.5点
EFTの平均語数	3点	2点	1点
合計	25.5点	17点	8.5点

3章3節の結果として得られたCEFR-J A2レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表わす説明変数の定量的な数値を示す表3-18を以下に再掲する。

表3-18 CEFR-J A2レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表わす説明変数の定量的な数値

	上位者	中位者	下位者
総語数	117以上	116～76	75以下
CLI	6.29以上	6.28～5.53	5.52以下
EFTの平均語数	8.35以上	8.34～7.06	7.05以下

表5-2と表3-18を一つの表にしてみる(表5-3)。

表5-3 習熟度の相違による言語的特徴を表わす説明変数の定量的な数値とその得点

	上位者	中位者	下位者
総語数	18点	12点	6点
	117以上	116～76	75以下
CLI	4.5点	3点	1.5点
	6.29以上	6.28～5.53	5.52以下
EFTの平均語数	3点	2点	1点
	8.35以上	8.34～7.06	7.05以下
合計	25.5点	17点	8.5点

表5-3がグループリックの基本となる評価の観点の基準値とその得点を示すことになる。

これらの分析の結果を基に、CEFR-J A2 レベルの自由英作文のルーブリックをデザインする。

5.2.5 CEFR-J A2 レベルの自由英作文のルーブリック

5.2.4 の分析結果を基に、表 5-4 に CEFR-J A2 レベルの自由英作文のルーブリックを提案する。

表 5-4 CEFR-J A2 レベルの自由英作文のルーブリック

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
情報性の豊かさ	情報量が十分に記述されており、内容がよくわかる。 得点：18 点（総語数 117 語以上）	伝えたい情報量がある程度記述されており、内容も一応理解できる。 得点：12 点（総語数 116 語～76 語）	伝えたい情報量が少なく、内容が理解しにくい。 得点：6 点（総語数 75 語以下）
語彙や文体の難易度	語彙や文体の難易度が高い。 得点：4.5 点（CLI:6.29 以上）	語彙や文体の難易度が普通である。 得点：3 点（CLI:6.28～5.53）	語彙や文体の難易度がやや低い。 得点：1.5 点（CLI:5.52 以下）
エラーが少なく、熟達した英文	エラーが少なく、十分に熟達した英文である。 得点：3 点（EFT の平均語数：8.35 以上）	ややエラーがあるものの、熟達した英文である。 得点：2 点（EFT の平均語数：8.34～7.06）	エラーがあり、あまり熟達した英文になっていない。 得点：1 点（EFT の平均語数：7.05 以下）

このルーブリックについて「5.1.2 本研究のルーブリックの新規性」で提示した 5 項目を基に、以下に説明する。

「(1)自由英作文の言語的特徴を重回帰分析により抽出した特定要因の評価指標を評価の観点とする」については、重回帰分析の結果、抽出された 3 説明変数である総語数（情報性の豊かさ）、CLI（語彙や文体の難易度）、EFT 平均語数（エラーが少なく、熟達した英文）を使用した。

「(2)特定要因の評価指標を評価の観点とした分析的評価とする」については、重回帰分析の結果、抽出された 3 説明変数による 3 項目による分析的評価が可能となる。

「(3)評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値する」については、評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値を基に、上・中・下位者の3水準に分けた。

「(4)評価の観点に対する評価基準（到達度合い）の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値（全体的評価に対する評価の観点の貢献度）の比率により設定する」については、評価の観点に対する評価基準（到達度合い）の得点は重回帰分析の各説明変数の標準化係数 β 値の比率を基に算出している。

「(5)評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、学習者や指導者に提供することができる」については、「評価を示す記述（例えば、「情報性の豊かさ」について「情報量が十分に記述されており、内容がよくわかる」）」の下部に定量的数値を記載した。「情報性の豊かさ」の上位者の定量的数値を117語以上と記載した。

「評価を示す記述」は、評価の観点の水準ごとに記載されるものであり、一般的に抽象的な表現となりがちである。本研究で提案しているCEFR-J A2レベルの自由英作文のルーブリックにおいても、「評価を示す記述」はいずれも抽象的な表現に留まっている。しかしながら、その下部に定量的な数値情報を提示することによって、「評価を示す記述」の欠点である抽象性を克服し、より具体的な評価に関する情報をルーブリックの使用者に提供することが可能となる。

最後に、公立高等学校で30年以上教壇に立つ、いわゆるベテランのA英語教諭にこのルーブリックを提示した際の、感想を以下に記述^{注33)}し、このルーブリックが高等学校の自由英作文指導の変革の起爆剤になりうる可能性を探ることとする。

このA先生は英語も堪能で、勤務する県の教育委員会によって選抜され、その県に所在する国立大学大学院教育学研究科へ県費によって派遣され、修士号を有する方である。さらに、常に英語の授業を工夫され、斬新なアイデアで高校生に英語の指導をされてこられた先生でもある。

その彼にして、このルーブリックを提示した時の当初の反応は懐疑的なものであった。つまり、自由英作文の指導は大変なのだという固定観念を持ってみえたのである。筆者は内心、大変な衝撃を受けていた。この先生ならルーブリックを見ただけで、すぐに高評価がいただけると過信していたからである。本当にこの3つの評価の観点を向上させれば、自由英作文の質が向上するのか疑問に感じてみえるようであった。自由英作文の指導の大変さを熟知され、困難な指導を数多く経験されてきているからであることはすぐに推察できた。本研究の1章4節の馬場（2010）や大井（2015）の指摘に直面することになった。「平林（筆者）さん、もうちょっと詳しく説明してくれる？私も理解できていないことも

注33) この記述の会話の部分は、A先生にこのルーブリックの感想を伺った時の内容を筆者の記憶に従って、記述したもので、一字一句を正確に表しているものではない。内容はほぼそのままである。

多そうだし。」と謙虚に言って下さり、筆者は、このルーブリックの評価の観点は、CEFR-J A2 レベルの自由英作文の言語的特徴を統計的に検証したことから説明し、どのようにして、このルーブリックをデザインしたかを解説させていただいた。

そうすると、少しずつ自由英作文の指導に対する固定観念が崩れ出し、特に、この評価の観点から指導のイメージをして下さいという解説あたりから、反応の変化が見られだした。「(筆者が) できるだけたくさん書かせるにはどうすればいいでしょうか(評価の観点の情報性の豊かさ)」のように質問すると「プレライティングの際に、アイデア出しの工夫をこらす・・・それから、中学校で習った簡単な語彙、作文でもすぐ使える語(まさに、本研究3章2節でCEFR-J A2 レベルの自由英作文で使われる語彙の90%程度は、そのような高頻度語彙)を使わせる・・・」と次々にアイデアを出された。「(筆者が) エラーが少なく、熟達した英文を書く、つまり、内容が伝わらないような英文のエラーを避けるようにし、簡易な複文を作るようにしてはどうでしょう。例えば、I think that / I believe that を付けるだけで、アーギュメントの英文が作れ、エラーの少ない長い複文の英文となりますよね。」という提案をすると、これを聞いたA先生は筆者の提案よりもずっと優れた指導のアイデアを出され始めた。このような会話を繰り返すうちに、A先生の自由英作文の指導に対する固定観念が変わり、このルーブリックの有用性を認識していただくことができ、「そういえば、関西のある大学のB教授が自由英作文は、まず、量を書かせることが大切であると言っていたのを公演で聞きました。このルーブリックを使って指導すれば、一定量の英文を生徒に書かせやすくすることは確かですね。」のように言われ、このルーブリックの有効性を認識していただくことができた。

このA先生の当初のルーブリックに対する反応が多く先生方が抱く考え方ではないかと思われる。このような評価の観点を生徒に教授してしまっているのかという疑念を抱かれるのがむしろ当然であろう。

しかし、自由英作文に関して高等学校で行われている実際の指導は、大井(2015)にあるように高等学校の先生方はどのように指導しているのか途方に暮れているような状態である。このルーブリックを活用すれば、A先生のようにその評価の観点から、様々な指導方法を自ら考案され、自由英作文の指導は困難という固定観念を打ち砕くことは可能なのである。つまり、ここで提案したCEFR-J A2 レベルのルーブリックは自由英作文の指導に革新をもたらすツールとなると言えよう。

5.2.6 結論

5章2節では、CEFR-J A2 レベルの学習者によって記述された自由英作文のルーブリックをデザインした(表5-4)。従来のルーブリックと異なるのは以下の点である。

- (1) 重回帰分析によって抽出されたCEFR-J A2 レベルの自由英作文の言語的特徴を表わす特定要因の評価指標である総語数(情報性の豊かさ)、CLI(語彙や文体の難易度)、EFT平均語数(エラーが少なく、熟達した英文)を評価の観点としていること。

- (2) 評価の観点である「情報性の豊かさ」、「語彙や文体の難易度」、「エラーが少なく、熟達した英文」に関して、分析的評価を採用していること。
- (3) 評価の観点である「情報性の豊かさ」、「語彙や文体の難易度」、「エラーが少なく、熟達した英文」に対する評価基準（到達度合い）を統計的に検証した 3 段階の定量的数値としていること。
- (4) 評価の観点である「情報性の豊かさ」、「語彙や文体の難易度」、「エラーが少なく、熟達した英文」に対する評価基準（到達度合い）の得点を各評価の観点を抽出した重回帰分析の標準化係数の比率により設定していること。
- (5) ルーブリックの中に、評価の観点の評価基準（到達度合い）ごとの定量的数値を掲載していること。

6 章では、この CEFR-J A2 レベルの自由英作文のルーブリックの使い方を説明する。また、自由英作文の全体的評価とこのルーブリックによって算出される得点の相関関係を調査し、ルーブリックの検証も行う。

5.3 CEFR-J B2 および C1 レベルの自由英作文ルーブリックの提案

5.3.1 はじめに

5章3節では、CEFR-J B2 および C1 レベルの自由英作文に関する言語的特徴を表わす特定要因の評価指標を評価の観点とする分析的評価に基づくルーブリックを提案する。

ルーブリックに関する先行研究については2章でまとめたが、CEFR-J B2 および C1 レベルを対象としたルーブリックはこれまで作成されていない。CEFR-J B2 および C1 レベルというのは日本人では上級の英語学習者であり、そのような学習者によって記述される自由英作文の質も高度である(資料1-2参照)。既存の自由英作文に関するルーブリックの最も一般的なものの内の1つの‘Independent WRITING Rubrics’では、大半が4点もしくは5点の範疇に入るほどである。そのため、英文の質に差のある場合でも、評価の峻別が必ずしも十分できているとは言えない状況である。また、‘Independent WRITING Rubrics’の評価の基準に一般的すぎる表現が使われており、その使用者に十分なフィードバックができていない状況にある。こうした状況を克服するためにも CEFR-J B2 および C1 レベルのルーブリックをデザインする意義は大きいと考える。

5.3.2 研究の目的

本研究の目的は CEFR-J B2 および C1 レベルの自由英作文に関する言語的特徴を表わす特定要因の評価指標を評価の観点とする分析的評価に基づくルーブリックをデザインすることである。そのルーブリックでは、評価の観点に対する評価基準(到達度合い)を統計的に検証した定量的数値を使用し、評価の観点に対する評価基準(到達度合い)の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値(全体的評価に対する評価の観点の貢献度)の比率により設定する。さらに、このルーブリックでは、評価を示す記述に加え、評価の観点の定量的な数値を掲載する。

5.3.3 研究方法

5.3.3.1 CEFR-J B2 および C1 レベルの調査参加者

調査参加者は4章1節と同一である(表4-2参照)。

5.3.3.2 手続き

- (1) Nishijima, et al (2007) の7つの質問に対する回答を検討し、ルーブリックの基本設計を行う。
- (2) 4章3節の結果として得られた自由英作文の言語的特徴を表わす特定要因の評価指標を評価の観点とする。
- (3) 評価の観点に対する評価基準(到達度合い)を4章3節で統計的に検証した定量的数値を使用する。
- (4) 評価の観点に対する評価基準(到達度合い)の得点を自由英作文の言語的特徴を抽出し

た重回帰分析の各説明変数の標準化係数 β 値の比率により設定する。
(5) 評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、一般的で抽象的な評価を示す記述を補完する。

5.3.4 CEFR-J B2 および C1 レベルの自由英作文のルーブリックのデザイン

5.3.4.1 Model 1 を利用した CEFR-J B2 および C1 レベルの自由英作文のルーブリック

表 5-1 に Nishijima, et al (2007) の質問項目とその回答を提示した。ここでは、5 章 2 節の CEFR-J A2 の自由英作文のルーブリックのデザインと同様に、Nishijima, et al (2007) の質問に対する回答をより詳細に示し、ルーブリックをデザインする。

4 章 3 節の CEFR-J の B2 および C1 レベルの自由英作文の言語的特徴を示す Model 1（表 4-12 参照）を利用したルーブリックをデザインする。Model 1 では自由英作文の言語的特徴を表わす説明変数として構成、内容、言語使用が抽出されている。

「(1) 誰がそのルーブリックを使用するのか」については、本研究の対象は CEFR-J B2 および C1 レベルの学習者とその指導者となる。

「(2) ルーブリックは広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するのに使われるのか、それとも、1 つのタスクに対する深い達成度を評価するのに使われるのか」については、広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するルーブリックとすることを意図する。この意図を達成するため、CEFR-J B2 および C1 レベルの自由英作文を統計的に実証した言語的特徴に関する結果を活用し、汎用化に耐えうるルーブリックのデザインすることとした。

「(3) ルーブリックは全体的評価とするのか、それとも、分析的評価とするのか」については、分析的評価を採用した。これは、分析的評価が全体的評価よりも信頼性の高いとされることや学習者に対して診断的フィードバックを提供するという利点があるためである (Hamp-Lyons, 1991a ; Jacobs et al, 1981)。

「(4) ルーブリックではどのような得点の付け方をするのか」については、4 章 3 節の CEFR-J B2 および C1 レベルの言語的特徴に関する実証研究の結果、説明変数となった 3 変数（構成、内容、言語使用）における重回帰分析の標準化係数（ β 値、表 4-12 参照）の比率と (7) の質問の回答の 3 水準を活用した。

「(5) 評価の観点には何を使用すべきか」については、4 章 3 節の CEFR-J B2 および C1 レベルの英語力の学習者によって記述された自由英作文の言語的特徴を表す特定要因の評価指標である構成、内容、言語使用を使用する。

「(6) 評価の観点はどのように定義すべきか」については、評価の観点が (5) の回答で提示した構成、内容、言語使用であり、この変数の算出方法は本研究において作成したものであるが（2 章 5 節 3 項参照）、用語は ESL Composition Profile でも使われているため、本研究のルーブリックでも、そのまま使うこととする。

「(7)その観点の水準をどのようにすべきか」については、4章3節で抽出された3説明変数（構成、内容、言語使用）に関して上・中・下位者の定量的な数値を算出しており、それをルーブリックの評価の観点の水準として使用する。

この7つの回答に加え、以下に得点算出に関わる部分(4)と(7)の回答を詳述する。(7)で3水準も用いることにしたので、各変数の上・中・下位者に対して、仮に3点、2点、1点を割り当てる。この仮の得点に(4)の回答による考え方を加味する。表4-12の「構成」、「内容」、「言語使用」の標準化係数（ β 値）がそれぞれ0.20、0.49、0.51となっており、ほぼ2：5：5の比率となっている。重回帰分析における β 値はその説明変数がどの程度の割合で従属変数（自由英作文の全体的評価）に貢献しているかを表す。この比率をルーブリックにおける得点の算出に活用する。

β 値の最も低い「構成」の上・中・下位者の3点、2点、1点を割り当て、「構成」、「内容」、「言語使用」の β 値の比率、2：5：5を掛け合わせることにする。このようにして、各説明変数の得点を算出したのが表5-5である。従って、この場合の最高点は18点で最小点は6点となる。

表5-5 各説明変数の得点

	上位者	中位者	下位者
構成	3点	2点	1点
内容	7.5点	5点	2.5点
言語使用	7.5点	5点	2.5点
合計	18点	12点	6点

4章3節の結果として得られたModel1から得られたCEFR-J B2およびC1レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値を示す表4-22を再掲する。

表4-22 Model1から得られたCEFR-J B2およびC1レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値

	上位者	中位者	下位者
構成	3.90以上	3.89～3.43	3.42以下
内容	4.90以上	4.89～3.48	3.47以下
言語使用	4.97以上	4.96～3.73	3.72以下

表5-5と表4-22を一つの表にしてみる（表5-6）。

表 5-6 習熟度の相違による言語的特徴を表わす説明変数の定量的な数値とその得点

	上位者	中位者	下位者
構成	3 点 3.90 以上	2 点 3.89～3.43	1 点 3.42 以下
内容	7.5 点 4.90 以上	5 点 4.89～3.48	2.5 点 3.47 以下
言語使用	7.5 点 4.97 以上	5 点 4.96～3.73	2.5 点 3.72 以下
合計	18 点	12 点	6 点

表 5-6 がルーブリックの基本となる評価の観点の基準値とその得点を示すことになる。

これらの分析の結果を基に、CEFR-J の B2 および C1 レベルの自由英作文のルーブリックをデザインする。

5.3.4.2 Model 2 を利用した CEFR-J B2 および C1 レベルの自由英作文のルーブリック

ここでは、5 章 2 節の CEFR-J A2 の自由英作文のルーブリックのデザインと同様に、Nishijima, et al (2007) の質問に対する回答をより詳細に示し、ルーブリックをデザインする。

4 章 3 節の CEFR-J の B2 および C1 レベルの自由英作文の言語的特徴を示す Model 2 (表 4-12 参照) を利用したルーブリックをデザインする。Model 2 では自由英作文の言語的特徴を表わす説明変数として言語使用、CLI が抽出されている。

質問項目(1)～(3)の回答については Model 1 と同一である。

「(1)誰がそのルーブリックを使用するのか」については、本研究の対象は CEFR-J B2 および C1 レベルの学習者とその指導者となる。

「(2)ルーブリックは広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するのに使われるのか、それとも、1つのタスクに対する深い達成度を評価するのに使われるのか」については、広範に渡るジャンルのタスクに対処しうる、広い意味合いでのライティングの達成度を評価するルーブリックとすることを意図する。この意図を達成するため、CEFR-J B2 および C1 レベルの自由英作文を統計的に実証した言語的特徴に関する結果を活用し、汎用化に耐えうるルーブリックのデザインすることとした。

「(3)ルーブリックは全体的評価とするのか、それとも、分析的評価とするのか」については、分析的評価を採用した。これは、分析的評価が全体的評価よりも信頼性の高いとされることや学習者に対して診断的フィードバックを提供するという利点があるためである (Hamp-Lyons, 1991a ; Jacobs et al, 1981)。

「(4)ルーブリックではどのような得点の付け方をするのか」については、5章3節の CEFR-J B2 および C1 レベルの言語的特徴に関する実証研究の結果、説明変数となった2変数（言語使用、CLI）における重回帰分析の標準化係数（ β 値、表 4-14 参照）の比率と(7)の質問の回答の3水準を活用した。

「(5)評価の観点には何を使用すべきか」については、4章3節の CEFR-J B2 および C1 レベルの英語力の学習者によって記述された自由英作文の言語的特徴を表す特定要因の評価指標である言語使用、CLI を使用する。

「(6)評価の観点はどのように定義すべきか」については、(5)の回答で提示した言語使用、CLI が評価の観点であり、前者については、この変数の算出方法は本研究において作成したものであるが（2章5節3項参照）、用語は ESL Composition Profile でも使われているため、本研究のルーブリックでも、そのまま使うこととする。一方、後者の CLI は5章2節の CEFR-J A2 レベルの自由英作文のルーブリックの評価の観点の定義「語彙や文体の難易度」を採用する。

「(7)その観点の水準をどのようにすべきか」については、4章3節で抽出された2説明変数（言語使用、CLI）に関して上・中・下位者の定量的な数値を算出しており、それをルーブリックの評価の観点の水準として使用する。

この7つの回答に加え、以下に得点算出に関わる部分(4)と(7)の回答を詳述する。(7)で3水準も用いることにしたので、各変数の上・中・下位者に対して、仮に3点、2点、1点を割り当てる。この仮の得点に(4)の回答による考え方を加味する。表 4-14 の「言語使用」、「CLI」の標準化係数（ β 値）がそれぞれ 0.59、0.33 ほぼ 2:1 の比率（正確には 1.8:1 であるが、より単純な比率を採用した）となっている。重回帰分析における β 値はその説明変数がどの程度の割合で従属変数(自由英作文の全体的評価)に貢献しているかを表す。この比率をルーブリックにおける得点の算出に活用する。

β 値の低い「CLI」の上・中・下位者の3点、2点、1点を割り当て、「言語使用」との比率2を掛け合わせることにする。このようにして、各説明変数の得点を算出したのが表 5-7 である。従って、この場合の最高点は9点で最小点は3点となる。

表 5-7 各説明変数の得点

	上位者	中位者	下位者
言語使用	6点	4点	2点
CLI	3点	2点	1点
合計	9点	6点	3点

4章3節の結果として得られた Model 2 から得られた CEFR-J B2 および C1 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値を示す

表 4-23 を再掲する。

表 4-23 Model 2 から得られた CEFR-J B2 および C1 レベルにおける自由英作文能力の習熟度の相違による言語的特徴を表す説明変数の定量的な数値

	上位者	中位者	下位者
言語使用	4.97 以上	4.96～3.73	3.72 以下
CLI	12.24 以上	12.23～10.83	10.82 以下

表 5-7 と表 4-23 を一つの表にしてみる (表 5-8)。

表 5-8 習熟度の相違による言語的特徴を表わす説明変数の定量的な数値とその得点

	上位者	中位者	下位者
言語使用	6 点 4.97 以上	4 点 4.96～3.73	2 点 3.72 以下
CLI	3 点 12.24 以上	2 点 12.23～10.83	1 点 10.82 以下
合計	9 点	6 点	3 点

表 5-8 がルーブリックの基本となる評価の観点の基準値とその得点を示すことになる。これらの分析の結果を基に、CEFR-J の B2 および C1 レベルの自由英作文のルーブリックをデザインする。

5.3.5 CEFR-J の B2 および C1 レベルの自由英作文のルーブリック

5.3.5.1 Model 1 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック

5.3.4.1 の分析結果を基に、表 5-9 に Model 1 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリックを提案する。

表 5-9 Model 1 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック

評価の観点	評価基準 (到達の度合い)		
	上位者	中位者	下位者
構成	序論で与えられたトピックを適切に把握し、本論に非常に適格につながる構成をしている。結	序論で与えられたトピックを把握し、本論につながる構成をしている。結論で本論と首尾一貫	序論で与えられたトピックを把握し、本論につながる構成を成していない。結論で本論と首尾

	論で本論と十分に首尾一貫したつながりを有している。 得点：3.0点 (構成の値：3.90以上)	したつながりを有している。 得点：2.0点 (構成の値：3.89～3.43)	一貫したつながりを有していない。 得点：1.0点 (構成の値：3.42以下)
内容	本論の主題部分の内容が非常に適切であり、その支持部分の内容も極めて適切である。 得点：7.5点 (内容の値：4.90以上)	本論の主題部分の内容が適切であり、その支持部分の内容も適切である。 得点：5.0点 (内容の値：4.89～3.48)	本論の主題部分の内容が必ずしも適切ではなく、その支持部分も適切ではない。 得点：2.5点 (内容の値：3.47以下)
言語使用	語彙や文法の誤りがほとんどみられない。 得点：7.5点 (言語使用の値：4.97以上)	語彙や文法の誤りがややみられる。 得点：5.0点 (言語使用の値：4.96～3.73)	語彙や文法の誤りがかなりみられる。 得点：2.5点 (言語使用の値：3.72以下)

このルーブリックについて「5.1.2 本研究のルーブリックの新規性」で提示した5項目を基に、以下に説明する。

「(1)自由英作文の言語的特徴を重回帰分析により抽出した特定要因の評価指標を評価の観点とする」については、重回帰分析の結果、抽出された3説明変数である構成、内容、言語使用を使用した。

「(2)特定要因の評価指標を評価の観点とした分析的評価とする」については、重回帰分析の結果、抽出された3説明変数による3項目による分析的評価が可能となる。

「(3)評価の観点に対する評価基準(到達度合い)を統計的に検証した定量的数値とする」については、評価の観点に対する評価基準(到達度合い)を統計的に検証した定量的数値を基に、上・中・下位者の3水準に分けた。

「(4)評価の観点に対する評価基準(到達度合い)の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値(全体的評価に対する評価の観点の貢献度)の比率により設定する」については、評価の観点に対する評価基準(到達度合い)の得点は重回帰分析の各説明変数の標準化係数 β 値の比率を基に算出している。

「(5)評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、学習者や指導者に提供することができる」については、「評価を示す記述」は、評価の観点の水準ごとに記載されるものであり、一般的に抽象的な表現となりがちである。本研究で提案している CEFR-J B2 および C1 レベルの自由英作文のルーブリックにおいても、「評価を示す記述」はいずれも抽象的な表現に留まっている。しかしながら、その下部に定量的な数値情報を提示することによって、「評価を示す記述」の欠点である抽象性を克服し、より具体的な評価に関する情報をルーブリックの使用者に提供することが可能となる。

5.3.5.2 Model 2 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック

5.3.4.2 の分析結果を基に、表 5-10 に Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを提案する。

表 5-10 Model 2 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
言語使用	語彙や文法の誤りがほとんどみられない。 得点：6 点 （言語使用の値：4.97 以上）	語彙や文法の誤りがややみられる。 得点：4 点 （言語使用の値：4.96～3.73）	語彙や文法の誤りがかなりみられる。 得点：2 点 （言語使用の値：3.72 以下）
語彙や文体の難易度	語彙や文体の難易度が高い。 得点：3 点 （CLI:12.24 以上）	語彙や文体の難易度が普通である。 得点：2 点 （CLI:12.23～10.83）	語彙や文体の難易度がやや低い。 得点：1 点 （CLI:10.82 以下）

このルーブリックについて「5.1.2 本研究のルーブリックの新規性」で提示した 5 項目を基に、以下に説明する。

「(1)自由英作文の言語的特徴を重回帰分析により抽出した特定要因の評価指標を評価の観点とする」については、重回帰分析の結果、抽出された 2 説明変数である言語使用、CLI を使用した。

「(2)特定要因の評価指標を評価の観点とした分析的評価とする」については、重回帰分析の結果、抽出された 3 説明変数による 2 項目による分析的評価が可能となる。

「(3)評価の観点に対する評価基準(到達度合い)を統計的に検証した定量的数値とする」

については、評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値を基に、上・中・下位者の3水準に分けた。

「(4)評価の観点に対する評価基準（到達度合い）の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値（全体的評価に対する評価の観点の貢献度）の比率により設定する」については、評価の観点に対する評価基準（到達度合い）の得点は重回帰分析の各説明変数の標準化係数 β 値の比率を基に算出している。

「(5)評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、学習者や指導者に提供することができる」については、「評価を示す記述」は、評価の観点の水準ごとに記載されるものであり、一般的に抽象的な表現となりがちである。本研究で提案している CEFR-J B2 および C1 レベルの自由英作文のルーブリックにおいても、「評価を示す記述」はいずれも抽象的な表現に留まっている。しかしながら、その下部に定量的な数値情報を提示することによって、「評価を示す記述」の欠点である抽象性を克服し、より具体的な評価に関する情報をルーブリックの使用者に提供することが可能となる。

5.3.6 結論

5章3節では、CEFR-J B2 および C1 レベルによって記述された自由英作文の2つのルーブリックをデザインした（表5-9および表5-10）。Model 1に基づく、CEFR-J B2 および C1 レベルの自由英作文のルーブリックと従来のルーブリックと異なるのは以下の点である。

- (1) 重回帰分析によって抽出された CEFR-J B2 および C1 レベルの自由英作文の言語的特徴を表わす特定要因の評価指標である構成、内容、言語使用を評価の観点としていること。
- (2) 評価の観点である構成、内容、言語使用に関して、分析的評価を採用していること。
- (3) 評価の観点である構成、内容、言語使用に対する評価基準（到達度合い）を統計的に検証した定量的数値としていること。
- (4) 評価の観点である構成、内容、言語使用に対する評価基準（到達度合い）の得点を各評価の観点を抽出した重回帰分析の標準化係数の比率により設定していること。
- (5) ルーブリックの中に、評価の観点の評価基準（到達度合い）ごとの定量的数値を掲載していること。

Model 2に基づく、CEFR-J B2 および C1 レベルの自由英作文のルーブリックと従来のルーブリックと異なるのは以下の点である。

- (1) 重回帰分析によって抽出された CEFR-J B2 および C1 レベルの自由英作文の言語的特徴を表わす特定要因の評価指標である言語使用、CLI を評価の観点としていること。
- (2) 評価の観点である言語使用、CLIに関して、分析的評価を採用していること。
- (3) 評価の観点である言語使用、CLIに対する評価基準（到達度合い）を統計的に検証した定量的数値としていること。
- (4) 評価の観点である言語使用、CLIに対する評価基準（到達度合い）の得点を各評価の観点を抽出した重回帰分析の標準化係数の比率により設定していること。

(5) ルーブリックの中に、評価の観点の評価基準（到達度合い）ごとの定量的数値を掲載していること。

6章では、この CEFR-J B2 および C1 レベルの自由英作文のルーブリックの使い方を説明する。また、自由英作文の全体的評価とこのルーブリックによって算出される得点の相関関係を調査し、ルーブリックの検証も行う。

5.4 まとめ

5章では、Nishijima, et al (2007) の7項目に沿って、本研究のルーブリックのデザインに関する基本的な考え方を示した。その考え方に基づき、本研究のルーブリックは以下の5点において既存のルーブリックと異なる新規性を有する。

- (1) 重回帰分析により抽出した自由英作文の言語的特徴を特定要因の評価指標を評価の観点とすること。
- (2) 特定要因の評価指標を評価の観点とした分析的評価とすること。
- (3) 評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値とすること。
- (4) 評価の観点に対する評価基準（到達度合い）の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値（全体的評価に対する評価の観点の貢献度）の比率により設定すること。
- (5) 評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、学習者や指導者に提供することができること。

このような新規性を有する CEFR-J A2 レベルの自由英作文のルーブリックは以下のようになった。

- (1) 重回帰分析により抽出した自由英作文の言語的特徴である総語数（情報性の豊かさ）、CLI（語彙や文体の難易度）、EFT 平均語数（エラーが少なく、熟達した英文）を評価の観点としている。
- (2) 評価の観点である「情報性の豊かさ」、「語彙や文体の難易度」、「エラーが少なく、熟達した英文」に関して、分析的評価を採用している。
- (3) 評価の観点である「情報性の豊かさ」、「語彙や文体の難易度」、「エラーが少なく、熟達した英文」に対する評価基準（到達度合い）を統計的に検証した上・中・下位者という3段階の定量的数値としている。
- (4) 評価の観点である「情報性の豊かさ」、「語彙や文体の難易度」、「エラーが少なく、熟達した英文」に対する評価基準（到達度合い）の得点を各評価の観点を抽出した重回帰分析の標準化係数の比率（6 : 1.5 : 1）により設定している。
- (5) 評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を（「情報性の豊かさ」は、上位者：総語数 117 語以上、中位者：総語数 117 語～76 語、下位者：総語数 75 語以下、「語彙や文体の難易度」は、上位者：CLI 6.29

以上、中位者：CLI 6.28～5.53、下位者：CLI 5.52 以下、「エラーが少なく、熟達した英文」は、上位者：EFT の平均語数 8.35 以上、中位者：EFT の平均語数 8.34～7.06、下位者：EFT の平均語数 7.05 以下、) 掲載し、学習者や指導者に提供している。

CEFR-J B2 および C1 レベルの自由英作文のルーブリックは Model 1 と Model 2 の 2 つの自由英作文の言語的特徴のモデルが構築されたため、2 つのルーブリックをデザインした。

Model 1 に基づくレ CEFR-J B2 および C1 ベルの自由英作文のルーブリックは以下のようになった。

- (1) 重回帰分析により抽出した自由英作文の言語的特徴である構成、内容、言語使用を評価の観点としている。
- (2) 評価の観点である構成、内容、言語使用に関して、分析的評価を採用している。
- (3) 評価の観点である構成、内容、言語使用に対する評価基準（到達度合い）を統計的に検証した上・中・下位者という 3 段階の定量的数値としている。
- (4) 評価の観点である構成、内容、言語使用に対する評価基準（到達度合い）の得点を各評価の観点を抽出した重回帰分析の標準化係数の比率（2：5：5）により設定している。
- (5) 評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を（構成は、上位者：3.0 以上、中位者：3.89～3.43、下位者：3.42 以下、内容は、上位者：4.90 以上、中位者：4.89～3.48、下位者：3.47 以下、言語使用は、上位者：4.97 以上、中位者：4.96～3.73、下位者：3.72 以下、) 掲載し、学習者や指導者に提供している。

Model 2 に基づくレ CEFR-J B2 および C1 ベルの自由英作文のルーブリックは以下のようになった。

- (1) 重回帰分析により抽出した自由英作文の言語的特徴である言語使用、CLI を評価の観点としている。
- (2) 評価の観点である言語使用、CLI に関して、分析的評価を採用している。
- (3) 評価の観点である言語使用、CLI に対する評価基準（到達度合い）を統計的に検証した上・中・下位者という 3 段階の定量的数値としている。
- (4) 評価の観点である言語使用、CLI に対する評価基準（到達度合い）の得点を各評価の観点を抽出した重回帰分析の標準化係数の比率（2：1）により設定している。
- (5) 評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を（言語使用は、上位者：4.97 以上、中位者：4.96～3.73、下位者：3.72 以下、CLI は、上位者：12.24 以上、中位者：12.23～10.83、下位者：10.82 以下）掲載し、学習者や指導者に提供している。

6. 提案した自由英作文のルーブリックの使い方と検証

6.1 CEFR-J A2 レベルの自由英作文ルーブリックの使い方

CEFR-J A2 レベルの自由英作文ルーブリックの使い方を説明するために、表 5-4 を以下に再掲する。

表 5-4 CEFR-J A2 レベルの自由英作文のルーブリック

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
情報性の豊かさ	情報量が十分に記述されており、内容がよくわかる。 得点：18点（総語数 117 語以上）	伝えたい情報量がある程度記述されており、内容も一応理解できる。 得点：12点（総語数 116 語～76 語）	伝えたい情報量が少なく、内容が理解しにくい。 得点：6点（総語数 75 語以下）
語彙や文体の難易度	語彙や文体の難易度が高い。 得点：4.5点（CLI:6.29 以上）	語彙や文体の難易度が普通である。 得点：3点（CLI:6.28～5.53）	語彙や文体の難易度がやや低い。 得点：1.5点（CLI:5.52 以下）
エラーが少なく、熟達した英文	エラーが少なく、十分に熟達した英文である。 得点：3点（EFT の平均語数：8.35 以上）	ややエラーがあるものの、熟達した英文である。 得点：2点（EFT の平均語数：8.34～7.06）	エラーがあり、あまり熟達した英文になっていない。 得点：1点（EFT の平均語数：7.05 以下）

CEFR-J A2 レベルの学習者によって記述された自由英作文の評価者は以下の手続きで、「情報性の豊かさ」を表わす総語数、「語彙や文体の難易度」を表わす CLI、「エラーが少なく、熟達した英文」を表わす EFT 平均語数を算出する。

- (1) 総語数については、自由英作文のテキストがデジタルデータ化してあれば、MS-ワードの「スペルチェックと文章構成」機能により、算出が可能となる。もちろん、「文字カウント」機能によっても総語数は算出が可能であるが、「スペルチェックと文章構成」を活用するとリーダビリティの指標の一つである FKG も算出できるので、CLI との比較もできることになる。
- (2) CLI は Word Level Checker (Someya, 2009) により算出することができる。
- (3) EFT 平均語数に関しては、筆者に知る限り、算出するソフトウェアは開発されていないようである。自由英作文のテキストがデジタル化してあれば、エラーの定義を

明確にし、MS-ワードの「文字カウント」機能を活用すれば、短時間で算出が可能である。本研究の基となる 106 個の自由英作文のテキストの総語数の平均は約 98 語、平均文数が 10 文程度（表 3-10 参照）であり、EFT 平均語数の算出に慣れない当初はやや手間取るかもしれないが、慣熟するにつれ短時間で算出できるようになる。

採点の例を以下に示す。例えば、総語数 120 語、CLI が 6.00、EFT 平均語数が 8.55 の自由英作文の場合を想定する。総語数は上位者に入り、18 点となり、CLI は中位者に入り、3 点、EFT 平均語数は上位者に入り、3 点となり、合計 24 点となる。このように、総語数、CLI、EFT 平均語数を算出すれば、表 5-4 の CEFR-J A2 レベルの自由英作文のルーブリックを参考にして、評価を行うだけとなる。極端な言い方をすれば、自由英作文を読まずに、この総語数、CLI、EFT 平均語数の数値だけを基に、自由英作文の評価ができることになる。

そこで、CEFR-J A2 レベルの自由英作文の全体的評価とこのルーブリックを基に算出した得点との相関関係を調査し、ルーブリックを検証することにする。

6.2 CEFR-J A2 レベルの自由英作文のルーブリックの検証

表 5-4 の CEFR-J A2 レベルの自由英作文のルーブリックを用いて、本研究 CEFR-J A2 レベルの 106 個の自由英作文のテキストで得点を算出する。その得点と 106 個の自由英作文の全体的評価との相関関係を調査することによってルーブリックの検証を行った。表 6-1 はデザインしたルーブリックを基に各観点の得点を算出した記述統計である。また、表 6-2 は、この 106 個の CEFR-J A2 レベルの自由英作文の全体的評価である WS のみの記述統計を表 3-1 から抜粋したものである。

表 6-1 デザインしたルーブリックによる得点の記述統計 (N = 106)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
情報性の豊かさ	12.00	4.38	6.00	18.00	12.00	.00	-.87
語彙や文体の難易度	2.94	1.46	1.50	4.50	3.00	.08	-2.00
エラーが少なく熟達した英文	2.03	.84	1.00	3.00	2.00	-.05	-1.60
合計	16.97	5.07	8.50	25.50	17.00	-.11	-.93

表 6-2 CEFR-J A2 レベルの 106 個の全体的評価 (WS) の記述統計 (N = 106)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
WS	160	99.72	19.00	50	130	80	0.26	-0.87

表 6-1 の合計と表 6-2 の WS の相関関係を調査してみた。結果として、両者のピアソンの相関係数は 0.76 ($p < .01$ 両側検定)となり、本研究で提案した CEFR-J A2 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には、

正の高い相関関係があることが判明した。

このループリックの得点と全体的評価による散布図を確認したところ、106 個の自由英作文の内、5 個が外れ値と考えられた。いずれも、ループリックの得点は低いが、全体的評価は高いという外れ値であった。これらの 5 個の自由英作文を読み直してみた。その結果、情報性の豊かさ（総語数）、語彙や文体の難易度（CLI）、エラーが少なく熟達した英文（EFT 平均語数）以外の要因がとして、内容自体の面白さ、興味深さの切り口が優れており、これらの 5 個の自由英作文の全体的評価を高くしていると考えられることが判明した。

確かに、この提案した CEFR-J A2 レベルの自由英作文のループリックの 3 つの評価の観点では、内容面での評価を成し得ないという欠点もないとは言えない。しかしながら、106 個の自由英作文の内、5 個（5%程度）というのは外れ値の範疇に入るものであり、ループリックによる得点は全体的評価を著しく損ねるものではないことは明らかである。

また、上述したループリックのデザインに使用した 106 個の自由英作文を記述した学習者とは異なる CEFR-J A2 レベル学習者によって記述された自由英作文 63 個を基に、検証を試みることにした。この 63 個の自由英作文の課題は「日本の学校では生徒 1 人に 1 台のコンピュータを用意すべきだ、という意見があります。この意見に賛成か反対かあなたの立場を明らかにし、あなたの経験や知識をもとに、その理由をできるだけたくさん英語で述べなさい」というものである。

表 6-3 は、検証に使用した自由英作文を記述した学習者の英語力に関する記述統計である。この学習者の英語力の測定に使用したテストは GTEC であり、800 点満点で、公式により TOEIC の得点に換算が可能である（3 章 1 節 4 項参照）。TOEIC の平均点は 325 点であり、ループリックのデザインに使用した 106 個の自由英作文を記述した学習者の平均点が 373 点であったことと比較すると、やや英語力は劣るものの、最小値 231 点～最大値 497 点であり、CEFR-J A2 レベル（TOEIC 225 点～550 点）全般をほぼ満たすと考えられる。この表 6-3 の中の WS が自由英作文の全体的評価となる。

表 6-3 検証に使用した自由英作文を記述した学習者の英語力の記述統計（ $N = 63$ ）

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
R S	320	162.54	36.28	95	242	147	0.41	-0.55
L S	320	179.63	26.74	135	241	106	0.46	-0.22
WS	160	87.94	16.48	50	130	80	0.19	1.51
総合点	800	430.11	66.66	339	595	256	0.80	-0.11
TOEIC								
換算点	990	325.32	69.33	231	497	266	0.80	-0.11

RS:リーディングセクション LS:リスニングセクション WS:ライティングセクション

次に、この検証に使用した 63 個の自由英作文に関して、デザインしたループリックの

評価の観点である総語数、CLI、EFT 平均語数の記述統計を表 6-4 に示す。

表 6-4 検証に使用した 63 個の自由英作文に関する総語数、CLI、EFT 平均語数の記述統計

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
総語数	94.81	34.22	33.00	191.00	158.00	0.76	0.26
C L I	6.60	1.87	3.10	10.80	7.70	0.30	-0.63
EFT 平均語数	8.35	2.30	4.83	14.33	9.50	0.74	0.07

この表 6-4 で示した総語数、CLI、EFT 平均語数に関して、デザインした CEFR-J A2 レベルの自由英作文のルーブリックを基に、各評価の観点の得点を算出し、その記述統計を表 6-5 に示した。

表 6-5 デザインしたルーブリックによる得点の記述統計 (N = 63)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
情報性の豊かさ	11.62	4.43	6.00	18.00	12.00	.10	-1.12
語彙や文体の難易度	3.31	1.38	1.50	4.50	3.00	-.43	-1.70
エラーが少なく熟達した英文	2.10	.89	1.00	3.00	2.00	-.19	-1.74
合計	17.02	4.65	8.50	25.50	17.00	-.18	-.79

この表 6-5 の合計と表 6-3 の WS の相関関係を調査してみた。結果として、ピアソンの相関係数は 0.75 ($p < .01$ 両側検定) となり、本研究で提案した CEFR-J A2 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には正の高い相関関係があることが判明した。

このルーブリックの得点と全体的評価による散布図を確認したところ、63 個の自由英作文の内、4 個が外れ値と考えられた。いずれも、ルーブリックの得点は高いが、全体的評価はやや低いという外れ値であった。これらの 4 個の自由英作文を読み直してみた。その結果、情報性の豊かさ(総語数)に優れており、ルーブリックの得点が高くなっているが、一方、内容自体の面白さに欠け、全体的評価を低くしていることが判明した。しかしながら、63 個の自由英作文の内、4 個(6%程度)というのは外れ値の範疇に入るものであり、ルーブリックによる得点は全体的評価を著しく損ねるものではないとことは明らかである。

これら結果から、本研究で提案した CEFR-J A2 レベルの自由英作文のルーブリックの評価は全体的評価に代用しうるものと言え、その有効性が検証された。

6.3 CEFR-J B2 および C1 レベルの自由英作文のルーブリックの使い方

6.3.1 Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの使い方

Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの使い方を説明するために、表 5-9 を以下に再掲する。

表 5-9 Model 1 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
構成	序論で与えられたトピックを適切に把握し、本論に非常に適格につながる構成をしている。結論で本論と十分に首尾一貫したつながりを有している。 得点：3.0 点 (構成の値：3.90 以上)	序論で与えられたトピックを把握し、本論につながる構成をしている。結論で本論と首尾一貫したつながりを有している。 得点：2.0 点 (構成の値：3.89～3.43)	序論で与えられたトピックを把握し、本論につながる構成を成していない。結論で本論と首尾一貫したつながりを有していない。 得点：1.0 点 (構成の値：3.42 以下)
内容	本論の主題部分の内容が非常に適切であり、その支持部分の内容も極めて適切である。 得点：7.5 点 (内容の値：4.90 以上)	本論の主題部分の内容が適切であり、その支持部分の内容も適切である。 得点：5.0 点 (内容の値：4.89～3.48)	本論の主題部分の内容が必ずしも適切ではなく、その支持部分も適切ではない。 得点：2.5 点 (内容の値：3.47 以下)
言語使用	語彙や文法の誤りがほとんどみられない。 得点：7.5 点 (言語使用の値：4.97 以上)	語彙や文法の誤りがややみられる。 得点：5.0 点 (言語使用の値：4.96～3.73)	語彙や文法の誤りがかなりみられる。 得点：2.5 点 (言語使用の値：3.72 以下)

この Model 1 に基づくルーブリックの評価の観点（自由英作文の言語的特徴）である「構

成」、「内容」、「言語使用」の算出方法を以下に簡潔に説明する（詳細は2章5節3項参照）。CEFR-J B2 および C1 レベルの自由英作文は、その英語力が TOEIC 785 点以上という高度なものであることを考慮して、そのライティング力も相当レベルが高いと考えた。そこで、その自由英作文は序論、本論、結論の形式で記述されていることを前提^{注34)}として、「構成」「内容」という特定要因の評価指標を作成した。

「構成」という視点からすると序論および結論という体裁を成していれば、本論も必然的に存在すると考え、序論および結論の部分を採点の対象とする変数を作成する。

「内容」に関しては、本論の部分を対象とし、それを主題部分と支持部分に分け、採点の対象とする変数を作成する。

また、「言語使用」に関しては、TOEIC の平均得点で 907 点である上級英語学習者のため、語彙や文法のローカルエラーまでも誤りとする厳格なものとし、エラーの重み付けはせず、誤りがあればすべてエラーとして算出する。

以下にその算出方法を記述する。「構成」は序論と結論の体裁が整っていることを前提であり、序論については、与えられたトピックを踏まえ、本論につながる構成をしているかどうかを2点満点（0.5点刻み）で採点し、結論については、本論と首尾一貫したつながりを有しているかどうかを2点満点（0.5点刻み）で採点する。従って、最高点は4点、最低点は0点となる特定要因の評価指標となる。

「内容」は本論を主題部分と支持部分に分け、主題部分の内容が適切かどうかを3点満点（0.5点刻み）で採点し、支持部分の内容が適切かどうかを3点満点（0.5点刻み）で採点する。従って、最高点は6点、最低点は0点となる特定要因の評価指標となる。

「言語使用」は語彙や文法の誤りに関して、エラーの数で5段階の得点を付与した。平均総語数が約230語、1文平均語数が約12語（表4-11参照）のため、5点は40語（3文程度）ごとに平均1つのエラーまで、4点は20語（1.5文程度）ごとに平均1つのエラーまで、3点は15語（1文程度）に平均1つのエラーまで、2点は10語ごとに平均1つのエラーまで、1点は10語に至るまでに平均最低1つのエラーとした。従って、最高点は5点、最低点は1点となる特定要因の評価指標となる。

このような評価を行うために、以下のような評価シート（表6-6～表6-11）を作成し、「構成」、「内容」、「言語使用」ごとの評価を行い、各評価の観点の得点を算出する。その際に、自由英作文が5パラグラフで構成されている場合は評価シート（表6-6～表6-8）を、4パラグラフで構成されている場合は評価シート（表6-9～表6-11）を使用する。通常、CEFR-J B2 および C1 レベルの上級英語学習者になると、30分で自由英作文を記述する場合、序論・本論・結論の3段構成で、序論で1パラグラフ、本論で2もしくは3パラグラフ、そして結論で1パラグラフの4ないし5パラグラフの構成になるというメタ知識を有していると考えられる。実際、65個の自由英作文のすべてが、序論・本論・結論の3段構成を

^{注34)} 対象とした65個の自由英作文すべてが、序論、本論、結論の形式で記述されていた。

成し、4ないし5パラグラフであった。但し、本論を3つのパラグラフで構成するのが一般的とされ、実際、65個中54個(83%)の自由英作文が5パラグラフ構成であった。

まず、5パラグラフ(本論が3パラグラフ)で構成されている自由英作文の各評価の観点の得点の算出方法を説明する。表6-6で「構成」に関する得点を算出する。序論で、与えられたトピックを踏まえて、本論につながる構成をしているかどうかを2点満点の5段階で評価し、表6-6の(ア)の部分に得点を記入する。結論で本論と首尾一貫しているかどうかを2点満点の5段階で評価し、表6-6の(イ)の部分に得点を記入する。構成の合計点の欄に(ア)と(イ)を加えた結果(ウ)を記入する。

表 6-6 「構成」に関する評価シート

構成	序論		
	与えられたトピックを踏まえ、うまく本論につながる構成をしている。	2点	(ア)
	与えられたトピックを踏まえ、本論につながる構成をしている。	1.5点	
	与えられたトピックを踏まえ、本論につながる構成がややできていない。	1.0点	
	与えられたトピックを踏まえ、本論につながる構成が殆どできていない。	0.5点	
	与えられたトピックを踏まえ、本論につながる構成が全くできていない。	0点	
	結論		
	本論と非常に首尾一貫したつながりがある。	2点	(イ)
	本論と首尾一貫したつながりがある	1.5点	
	本論と首尾一貫したつながりにやや問題がある。	1.0点	
	本論と首尾一貫したつながりが殆どない。	0.5点	
本論と首尾一貫したつながりが全くない。	0点		
構成の合計点 = (ア) + (イ) = (ウ)			

表 6-7 で「内容」に関する得点を算出する。1つめの本論(本論①)の主題部分の内容が適切であるかどうかを3段階で評価する。その1つめの本論の主題部分に対する支持部分の内容が適切であるかどうかを3段階で評価する。この主題部分の得点と支持部分の得点を合計し、その結果を表6-7の(エ)に記入する。2つめの本論(本論②)と3つめの本論(本論③)でも同様の手続きを行い、表6-7の(オ)および(カ)に結果を記入する。内容の合計点の欄に(エ)、(オ)、(カ)を加えた結果(キ)を記入する。

表 6-7 本論が 3 パラグラフの場合の「内容」に関する評価シート

内容	本論①の主題部分		本論①の支持部分		本論① の得点
	主題部分の内容が適切である。	1 点	支持部分の内容が適切である。	1 点	(エ)
	主題部分の内容がやや不十分である。	0.5 点	支持部分の内容がやや不十分である。	0.5 点	
	主題部分の内容が不十分である。	0 点	支持部分の内容が不十分である。	0 点	
	本論②の主題部分		本論②の支持部分		本論② の得点
	主題部分の内容が適切である。	1 点	支持部分の内容が適切である。	1 点	(オ)
	主題部分の内容がやや不十分である。	0.5 点	支持部分の内容がやや不十分である。	0.5 点	
	主題部分の内容が不十分である。	0 点	支持部分の内容が不十分である。	0 点	
	本論③の主題部分		本論③の支持部分		本論③ の得点
	主題部分の内容が適切である。	1 点	支持部分の内容が適切である。	1 点	(カ)
	主題部分の内容がやや不十分である。	0.5 点	支持部分の内容がやや不十分である。	0.5 点	
	主題部分の内容が不十分である。	0 点	支持部分の内容が不十分である。	0 点	
内容の合計点 = (エ) + (オ) + (カ) = (キ)					

表 6-8 で「言語使用」に関する得点を算出する。自由英作文での語彙や文法の誤りに関するエラーの数を算出し、そのエラー数で自由英作文の総語数を除する。例えば、総語数 280 語で語彙や文法に関するエラーが 6 個の場合、 $280 \text{ 語} \div 6 = 47$ となり、言語使用の得点は 5 点となる。そのように算出した結果を (ク) に記入する。

表 6-8 「言語使用」に関する評価シート

言語 使用	対象とする自由英作文のテキスト内 40 語以上で平均 1 つのエ ラー	5 点	(ク)
	対象とする自由英作文のテキスト内 20 語～39 語で平均 1 つ のエラー	4 点	
	対象とする自由英作文のテキスト内 15 語～19 語で平均 1 つ のエラー	3 点	
	対象とする自由英作文のテキスト内 10 語～14 語で平均 1 つ のエラー	2 点	
	対象とする自由英作文のテキスト内 10 語未満で平均 1 つのエ ラー	1 点	

表 6-6～表 6-8 のワークシート^{注35)}によって「構成」、「内容」、「言語使用」の得点を算出する。例えば、「構成 (ア) : 2 点」、「構成 (イ) : 1.5 点」、「内容 (エ) : 2 点」、「内容 (オ) : 1.5 点」、「内容 (カ) : 1.5 点」、「内容 (キ) : 5.0 点」「言語使用 (ク) : 5 点」の自由英作文を想定する。

この得点と表 5-9 の Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックをもとに、ルーブリックにおける評価点を算出する。この場合、「構成」は 3.5 点という値となるため、「構成」の評価基準（到達の度合い）の中位者（「構成」の基準値：3.90～3.42）に入り、ルーブリックの得点は 2.0 点となる。「内容」は 5.0 点という値となるため、「内容」の評価基準（到達の度合い）の上位者（「内容」の基準値：4.90 以上）に入り、ルーブリックの得点は 7.5 点となる。「言語使用」は 5.0 という値となるため、「言語使用」の評価基準（到達の度合い）の上位者（「言語使用」の基準値：4.97 以上）に入り、ルーブリックの得点は 7.5 点となる。従って、この自由英作文のルーブリックを基に算出した得点は 17 点（18 点満点）となる。

次に、4 パラグラフ（本論が 2 パラグラフ）で構成されている自由英作文の各評価の観点の得点の算出方法を説明する。

表 6-9 で「構成」に関する得点を算出する。序論で、与えられたトピックを踏まえて、本論につながる構成をしているかどうかを 2 点満点の 5 段階で評価し、表 6-9 の (ケ) の部分に得点を記入する。結論で本論と首尾一貫しているかどうかを 2 点満点の 5 段階で評価し、表 6-9 の (コ) の部分に得点を記入する。構成の合計点の欄に (ケ) と (コ) を加えた結果 (サ) を記入する。

^{注35)} これらのワークシートをエクセルのような表計算ソフトで作成しておくこと、得点の算出は容易になる。

表 6-9 「構成」に関する評価シート

構成	序論		
	与えられたトピックを踏まえ、うまく本論につながる構成をしている。	2点	(ケ)
	与えられたトピックを踏まえ、本論につながる構成をしている。	1.5点	
	与えられたトピックを踏まえ、本論につながる構成がややできていない。	1.0点	
	与えられたトピックを踏まえ、本論につながる構成が殆どできていない。	0.5点	
	与えられたトピックを踏まえ、本論につながる構成が全くできていない。	0点	
	結論		
	本論と非常に首尾一貫したつながりがある。	2点	(コ)
	本論と首尾一貫したつながりがある	1.5点	
	本論と首尾一貫したつながりにやや問題がある。	1.0点	
	本論と首尾一貫したつながりが殆どない。	0.5点	
本論と首尾一貫したつながりが全くない。	0点		
構成の合計点 = (ケ) + (コ) = (サ)			

表 6-10 で「内容」に関する得点を算出する。1 つめの本論（本論①）の主題部分の内容が適切であるかどうかを 3 段階で評価する。その 1 つめの本論の主題部分に対する支持部分の内容が適切であるかどうかを 3 段階で評価する。この主題部分の得点と支持部分の得点を合計し、その結果を表 6-10 の（シ）に記入する。2 つめの本論（本論②）も同様の手続きを行い、表 6-10 の（ス）に結果を記入する。内容の合計点の欄に（シ）と（ス）を加えた値に 1.5 を乗じて、その値を（セ）に記入する。

表 6-10 本論が 2 パラグラフの場合の「内容」に関する評価シート

内容	本論①の主題部分		本論①の支持部分		本論① の得点
	主題部分の内容が適切である。	1 点	支持部分の内容が適切である。	1 点	(シ)
	主題部分の内容がやや不十分である。	0.5 点	支持部分の内容がやや不十分である。	0.5 点	
	主題部分の内容が不十分である。	0 点	支持部分の内容が不十分である。	0 点	
	本論②の主題部分		本論②の支持部分		本論② の得点
	主題部分の内容が適切である。	1 点	支持部分の内容が適切である。	1 点	(ス)
	主題部分の内容がやや不十分である。	0.5 点	支持部分の内容がやや不十分である。	0.5 点	
	主題部分の内容が不十分である。	0 点	支持部分の内容が不十分である。	0 点	
	内容の得点 {(シ) + (ス)} × 1.5 = (セ)				

表 6-11 で「言語使用」に関する得点を算出する。自由英作文での語彙や文法の誤りに関するエラーの数を算出し、そのエラー数で自由英作文の総語数を除する。例えば、総語数 250 語で語彙や文法に関するエラーが 15 個の場合、 $250 \text{ 語} \div 15 \div 17$ となり、言語使用の得点は 3 点となる。そのように算出した結果を (ソ) に記入する。

表 6-11 「言語使用」に関する評価シート

言語 使用	対象とする自由英作文のテキスト内 40 語以上で平均 1 つのエラー	5 点	(ソ)
	対象とする自由英作文のテキスト内 20 語～39 語で平均 1 つのエラー	4 点	
	対象とする自由英作文のテキスト内 15 語～19 語で平均 1 つのエラー	3 点	
	対象とする自由英作文のテキスト内 10 語～14 語で平均 1 つのエラー	2 点	
	対象とする自由英作文のテキスト内 10 語未満で平均 1 つのエラー	1 点	

表 6-9～表 6-11 のワークシート^{注36)} によって「構成」、「内容」、「言語使用」の得点を算出する。例えば、「構成 (ケ) : 2.0 点」、「構成 (コ) : 1.5 点」、「内容 (シ) : 1.5 点」、「内容 (ス) : 1.5 点」、「内容 (セ) : 4.5 点」、「言語使用 (ソ) : 3 点」の自由英作文を想定する。

この得点と表 5-9 の Model 1 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリックをもとに、ルーブリックにおける評価点を算出する。この場合、「構成」は 3.5 点という値となるため、「構成」の評価基準（到達の度合い）の中位者（「構成」の基準値：3.90～3.42）に入り、ルーブリックの得点は 2.0 点となる。「内容」は 4.5 点という値となるため、「内容」の評価基準（到達の度合い）の中位者（「内容」の基準値：4.89～3.47）に入り、ルーブリックの得点は 5.0 点となる。「言語使用」は 3.0 という値となるため、「言語使用」の評価基準（到達の度合い）の下位者（「言語使用」の基準値：3.72 以下）に入り、ルーブリックの得点は 2.5 点となる。従って、この自由英作文のルーブリックを基に算出した得点は 9.5 点（18 点満点）となる。

6 章 4 節において CEFR-J B2 および C1 レベルの自由英作文の全体的評価とこのルーブリックを基に算出した得点との相関関係を調査し、ルーブリックを検証することにする。

^{注36)} これらのワークシートをエクセルのような表計算ソフトで作成しておく、得点の算出は容易になる。

6.3.2 Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの使い方

Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの使い方を説明するために、表 5-10 を以下に再掲する。

表 5-10 Model 2 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
言語使用	語彙や文法の誤りがほとんどみられない。 得点：6点 （言語使用の値：4.97 以上）	語彙や文法の誤りがややみられる。 得点：4点 （言語使用の値：4.96～3.73）	語彙や文法の誤りがかなりみられる。 得点：2点 （言語使用の値：3.72 以下）
語彙や文体の難易度	語彙や文体の難易度が高い。 得点：3点 （CLI:12.24 以上）	語彙や文体の難易度が普通である。 得点：2点 （CLI:12.23～10.83）	語彙や文体の難易度がやや低い。 得点：1点 （CLI:10.82 以下）

このルーブリックを使用するために、表 6-12 で「言語使用」に関する得点を算出する。自由英作文での語彙や文法の誤りに関するエラーの数を算出し、そのエラー数で自由英作文の総語数を除する。例えば、総語数 310 語で語彙や文法に関するエラーが 14 個の場合、 $310 \text{ 語} \div 14 = 22$ となり、「言語使用」の得点は 4 点となる。そのように算出した結果を(タ)に記入する。CLI は Word Level Checker (Someya, 2009) により算出することができる。そして、CLI が 11.5 となったと想定する。この場合、「言語使用」の値が 4 点のため、表 5-10 の Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの中位者（言語使用の値：4.96～3.73）となり、「言語使用」のルーブリックの得点は 4 点に、CLI の値が 11.5 のため、「語彙や文体の難易度」の中位者（CLI:12.23～10.83）となり、ルーブリックの得点は 2 点となる。従って、自由英作文の Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に算出した得点は 6 点（9 点満点）となる。

表 6-12 「言語使用」に関する評価シート

言語 使用	対象とする自由英作文のテキスト内 40 語以上で平均 1 つのエラー	5 点	(タ)
	対象とする自由英作文のテキスト内 20 語～39 語で平均 1 つのエラー	4 点	
	対象とする自由英作文のテキスト内 15 語～19 語で平均 1 つのエラー	3 点	
	対象とする自由英作文のテキスト内 10 語～14 語で平均 1 つのエラー	2 点	
	対象とする自由英作文のテキスト内 10 語未満で平均 1 つのエラー	1 点	

6.4 CEFR-J B2 および C1 レベルの自由英作文のルーブリックの検証

6.4.1 Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの検証

表 5-9 の Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを用いて、本研究 CEFR-J B2 および C1 レベルの 65 個の自由英作文のテキストで得点を算出する。その得点と 65 個の自由英作文の全体的評価との相関関係を調査することによってルーブリックの検証を行った。表 6-13 はデザインしたルーブリックを基に各観点の得点を算出した記述統計である。また、表 4-3 は、この 65 個の CEFR-J B2 および C1 レベルの自由英作文の全体的評価の記述統計を 4 章 1 節より再掲したものである。

表 6-13 デザインしたルーブリックによる得点の記述統計 ($N = 65$)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
構成	2.51	0.87	1.00	3.00	2.00	-1.21	-0.57
内容	5.23	1.91	2.50	7.50	5.00	-1.26	-0.16
言語使用	5.96	2.06	2.50	7.50	5.00	-0.83	-1.01
合計	13.70	3.32	8.00	18.00	10.00	-0.11	-1.35

表 4-3 自由英作文の全体的評価 ($N = 65$)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
全体的評価 (HE)	5.00	4.03	0.65	2.00	5.00	3.00	-0.26	0.04

表 6-13 の合計と表 4-3 の全体的評価の相関関係を調査してみた。結果として、両者のピアソンの相関係数は 0.83 ($p < .01$ 両側検定)となり、本研究で提案した Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には、正の高い相関関係がることが判明した。

このルーブリックの得点と全体的評価による散布図を確認したところ、65 個の自由英作文の内、3 個が外れ値と考えられた。いずれも、ルーブリックの得点は低いが、全体的評価は高いという外れ値であった。これらの 3 個の自由英作文を読み直してみた。その結果、構成、内容、言語使用以外の要因として、内容面での洞察力が優れており、これらの 3 個の自由英作文の全体的評価を高くしていると考えられる。但し、この 3 個の自由英作文には記述にやや問題があるようで、このルーブリックの評価の観点である内容の本論における主題と指示部分の適切性(首尾一貫性)が必ずしも十分でなかったと考えられる。

しかしながら、65 個の自由英作文の内、3 個 (5%程度) という外れ値の範疇に入るものであり、ルーブリックによる得点は全体的評価を著しく損ねるものではないことは明らかである。

また、上述したルーブリックのデザインに使用した 65 個の自由英作文を記述した学習

者とは異なる CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文 29 個を基に、検証を試みることにした。この 29 個の自由英作文の課題は ‘Some people think that universities should provide graduates with the knowledge and skills needed in the workplace. Others believe that the true function of a university should be to give access to knowledge for its own sake, regardless of whether the course is useful to an employer. What, in your opinion, should be the main function of a university?’ というものである。

表 6-14 は、検証に使用した自由英作文を記述した学習者の英語力に関する記述統計である。TOEIC の平均点は 861 点で、ループリックをデザインに使用した自由英作文を記述した学習者の TOEIC の平均点の 907 点と比較すると、やや英語力は劣るものの、上級英語学習者である。

表 6-14 検証に使用した自由英作文を記述した学習者の英語力の記述統計 ($N = 29$)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
TOEIC R	495	409.14	37.63	330	480	150	-0.23	-0.43
TOEIC L	495	451.90	23.24	410	495	85	0.16	-0.56
合計	990	861.38	55.88	785	965	180	0.22	-1.11

TOEIC R: リーディングセクション TOEIC L: リスニングセクション

表 6-15 は、この上級英語学習者によって記述された自由英作文の全体的評価の記述統計である。この評価は、ループリックをデザインした 65 個の自由英作文と同様の手法により印象度による全体的評価で採点した。この 29 個の自由英作文の全体的評価の平均値は 3.45 となった。この数値は、ループリックをデザインした 65 個の自由英作文における全体的評価の平均値の 4.03 と比較するとやや劣ることになった。これは検証に使用した自由英作文を記述した学習者の TOEIC の平均値が 861 点で、ループリックをデザインした学習者の TOEIC の平均値よりも 50 点程度、英語力が低いことを反映していると考えられる。

表 6-15 検証に使用した自由英作文の全体的評価 ($N = 29$)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
全体的評価 (HE)	5.00	3.45	0.51	2.00	5.00	3.00	1.25	1.58

次に、この検証に使用した 29 個の自由英作文に関して、デザインした Model 1 の CEFR-J B2 および C1 レベルの自由英作文のループリックの評価の観点である構成、内容、言語使用の記述統計を表 6-16 に示す。

表 6-16 検証に使用した自由英作文に関する Model 1 の評価の観点の記述統計 (N = 29)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
構成	2.90	0.94	2.00	4.00	2.00	0.22	-1.91
内容	4.10	0.79	3.00	6.00	3.00	0.27	-0.39
言語使用	3.03	1.24	1.00	5.00	4.00	0.05	-0.98

この表 6-16 で示した構成、内容、言語使用に関して、デザインした Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に、各評価の観点の得点を算出し、その記述統計を表 6-17 に示した。

表 6-17 検証に使用した自由英作文に関するルーブリックの得点の記述統計 (N = 29)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
構成	1.76	0.99	1.00	3.00	2.00	0.53	-1.86
内容	5.00	1.89	2.50	7.50	5.00	0.00	-1.18
言語使用	3.79	1.84	2.50	7.50	5.00	1.08	-0.21
合計	10.55	3.34	6.00	18.00	12.00	0.49	-0.63

この表 6-17 の合計と表 6-15 の全体的評価の相関関係を調査してみた。結果としてピアソンの相関係数は 0.81 ($p < .01$ 両側検定) となり、本研究で提案した CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には正の高い相関関係があることが判明した。

このルーブリックの得点と全体的評価による散布図を確認したところ、29 個の自由英作文の内、3 個が外れ値と考えられた。いずれも、ルーブリックの得点は高いが、全体的評価は低いという外れ値であった。これらの 3 個の自由英作文を読み直してみた。その結果、内容の面白みは欠けるが、構成と内容における首尾一貫性に優れ、文法や語彙に関するエラーも少ないため、ルーブリックの得点は高いが、全体的評価は低くなっていることが判明した。

しかしながら、29 個の自由英作文の内 3 個は外れ値の範疇に入るものであり、ルーブリックによる得点は全体的評価を著しく損ねるものではないと考える。

これらの検証結果から、本研究で提案した Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの評価は全体的評価に代用が可能であり、その有効性が検証された。

6.4.2 Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの検証

表 5-10 の Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを用いて、本研究 CEFR-J B2 および C1 レベルの 65 個の自由英作文のテキストで得点を算出する。そ

の得点と 65 個の自由英作文の全体的評価との相関関係を調査することによってルーブリックの検証を行った。表 6-18 はデザインしたルーブリックを基に各観点の得点を算出した記述統計である。また、表 4-3 は、この 65 個の CEFR-J B2 および C1 レベルの自由英作文の全体的評価の記述統計を 4 章 1 節より再掲した。

表 6-18 デザインしたルーブリックによる得点の記述統計 (N = 65)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
言語使用	4.77	1.65	2.00	6.00	4.00	-.83	-1.01
CLI	1.89	.81	1.00	3.00	2.00	.20	-1.45
合計	6.66	2.03	3.00	9.00	6.00	-.45	-1.10

*CLI はルーブリックの評価の観点としては「語彙や文体の難易度」

表 4-3 自由英作文の全体的評価 (N = 65)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
全体的評価 (HE)	5.00	4.03	0.65	2.00	5.00	3.00	-0.26	0.04

表 6-18 の合計と表 4-3 の全体的評価の相関関係を調査してみた。結果として、両者のピアソンの相関係数は 0.70 ($p < .01$ 両側検定)となり、本研究で提案した Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には、正の高い相関関係があることが判明した。ピアソンの相関係数が 0.70 となり、Model 1 の場合のピアソンの相関係数 0.83 より下回るのは Model 1 の決定係数 R^2 が 0.73 であり、Model 2 の決定係数 R^2 が 0.54 という差に起因すると考えられる。

このルーブリックの得点と全体的評価による散布図を確認したところ、65 個の自由英作文の内、4 個が外れ値と考えられた。2 つは、ルーブリックの得点は低い、全体的評価は高いという外れ値であった。残りの 2 つは、ルーブリックの得点は高い、全体的評価は低いという外れ値であった。

これらの 4 個の自由英作文を読み直してみた。その結果、前者のルーブリックの得点は低い、全体的評価は高い自由英作文では、内容面での洞察力が優れており、これらの 2 個の自由英作文の全体的評価を高くしていると考えられる。この 2 つの内の 1 つは、Model 1 によるルーブリックにおいても外れ値となった英文である。後者のルーブリックの得点は高い、全体的評価は低い自由英作文では、内容自体は面白みに欠けるが、記述においてエラーが少なく、このルーブリックの評価の観点である言語使用の得点が高い。

このように、提案したルーブリックの言語使用と語彙や文体の難易度 (CLI) では自由英作文の内容面の評価を成し得ないという欠点もないとは言えない。しかしながら、65 個の自由英作文の内、4 個 (6%程度) という外れ値の範疇に入るものであり、ルーブリック

による得点は全体的評価を著しく損ねるものではないことは明らかである。

また、上述したルーブリックのデザインに使用した 65 個の自由英作文を記述した学習者とは異なる CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文 29 個 (Model 1 の検証にも使用した自由英作文) を基に検証を試みることにした。この 29 個の自由英作文の全体的評価の記述統計である表 6-15 を以下に再掲する。

表 6-15 検証に使用した自由英作文の全体的評価 ($N = 29$)

	満点	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
全体的評価 (HE)	5.00	3.45	0.51	2.00	5.00	3.00	1.25	1.58

次に、この検証に使用した 29 個の自由英作文に関して、デザインした Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの評価の観点である言語使用、CLI の記述統計を表 6-19 に示す。

表 6-19 検証に使用した自由英作文に関する Model 2 の評価の観点の記述統計 ($N = 29$)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
言語使用	3.03	1.24	1.00	5.00	4.00	0.05	-0.98
CLI	14.60	1.91	7.60	19.00	7.60	0.24	-0.55

この表 6-19 で示した言語使用、CLI に関して、デザインした Model 1 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に、各評価の観点の得点を算出し、その記述統計を表 6-20 に示した。

表 6-20 検証に使用した自由英作文に関するルーブリックの得点の記述統計 ($N = 29$)

	平均値	標準偏差	最小値	最大値	範囲	歪度	尖度
言語使用	3.03	1.48	2.00	6.00	4.00	1.08	-0.21
CLI	2.93	0.26	2.00	3.00	1.00	-3.59	11.70
合計	5.97	1.50	4.00	9.00	5.00	1.09	-0.06

*CLI はルーブリックの評価の観点としては「語彙や文体の難易度」

この表 6-20 の合計と表 6-15 の全体的評価の相関関係を調査してみた。結果としてピアソンの相関係数は 0.70 ($p < .01$ 両側検定) となり、本研究で提案した CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には正の高い相関関係があることが判明した。

このルーブリックの得点と全体的評価による散布図を確認したところ、29 個の自由英作

文の内、4 個が外れ値と考えられた。いずれも、ルーブリックの得点は高いが、全体的評価は低いという外れ値であった。これらの4 個の自由英作文を読み直してみた。その結果、内容の面白みは欠けるが、文法や語彙に関するエラーが少ないため、ルーブリックの得点は高いが、全体的評価は低くなっていることが判明した。

しかしながら、29 個の自由英作文の内、4 個は外れ値の範疇に入るものであり、ルーブリックによる得点は全体的評価を著しく損ねるものではないと考える。

これらの検証結果から、本研究で提案した Model 2 の CEFR-J B2 および C1 レベルの自由英作文のルーブリックの評価は全体的評価に代用が可能であり、その有効性が検証された。

6.5 まとめ

6 章では、提案したルーブリックの使い方とその有効性を検証した。CEFR-J A2 レベルの自由英作文のルーブリックについては、評価の観点となる特定要因の評価指標である総語数、CLI、EFT の平均語数の算出方法を解説した。これらの3 指標の数値が算出できれば、提案したルーブリックを基に、各評価の観点について評価対象とする自由英作文がどの評価基準（到達の度合い）に該当するかを判断し、その評価基準のルーブリックの得点から自由英作文の評価が可能となる。

このルーブリックを基に、ルーブリックのデザインに使用した 106 個の自由英作文の得点を算出し、その全体的評価との相関関係を調査することによって、ルーブリックの検証を試みた。結果として、ピアソンの相関係数は 0.76 ($p < .01$ 両側検定) となり、本研究で提案した CEFR-J A2 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には、正の高い相関関係があった。また、ルーブリックのデザインに使用したのものとは異なる自由英作文 63 個に関して、このルーブリックを基に、得点を算出し、その全体的評価との相関関係を調査してみた。結果として、ピアソンの相関係数は 0.75 ($p < .01$ 両側検定) となり、両者には正の高い相関関係があることも判明した。これらの結果により、提案した CEFR-J A2 レベルの自由英作文のルーブリックの有効性が検証された。

CEFR-J B2 レベルおよび C1 の自由英作文のルーブリックについては、評価の観点となる構成、内容、言語使用の特定要因の評価指標の算出の方法を解説した。Model 1 では、これらの3 指標の数値を算出し、Model 2 では、言語使用については Model 1 と同様に算出し、CLI は CEFR-J A2 レベルにおける場合と同様の手法で算出する。この数値が算出されれば、提案したルーブリックを基に、各評価の観点について評価対象とする自由英作文がどの評価基準（到達の度合い）に該当するかを判断し、その評価基準のルーブリックの得点から自由英作文の評価が可能となる。

Model 1 のルーブリックを基に、ルーブリックのデザインに使用した 65 個の自由英作文の得点を算出し、その全体的評価との相関関係を調査することによって、提案したルーブリックの検証を行った。結果として、ピアソンの相関係数は 0.83 ($p < .01$ 両側検定) と

なり、本研究で提案した CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には、正の高い相関関係があった。また、ルーブリックのデザインに使用したものとは異なる自由英作文 29 個に関して、このルーブリックを基に、得点を算出し、その全体的評価との相関関係を調査してみた。結果として、ピアソンの相関係数は 0.81 ($p < .01$ 両側検定) となり、両者には正の高い相関関係があることも判明した。これらの結果により、提案した Model11 に基づく CEFR-J B2 および C1 レベルの自由英作文のルーブリックの有効性が検証された。

Model 2 のルーブリックを基に、ルーブリックのデザインに使用した 65 個の自由英作文の得点を算出し、その全体的評価との相関関係を調査することによって、提案したルーブリックの検証を行った。結果として、ピアソンの相関係数は 0.70 ($p < .01$ 両側検定) となり、本研究で提案した CEFR-J B2 および C1 レベルの自由英作文のルーブリックを基に算出した得点とその自由英作文の全体的評価との間には、正の高い相関関係があった。また、ルーブリックのデザインに使用したものとは異なる自由英作文 29 個に関して、このルーブリックを基に、得点を算出し、その全体的評価との相関関係を調査してみた。結果として、ピアソンの相関係数は 0.70 ($p < .01$ 両側検定) となり、両者には正の高い相関関係があることも判明した。これらの結果により、提案した Model2 に基づく CEFR-J B2 および C1 レベルの自由英作文のルーブリックの有効性が検証された。

7. 結論

7.1 本研究の結論

本論文は日本における自由英作文の教育の普及を阻害する大きな原因の1つである評価基準の指針であるルーブリックをデザインすることを目的に研究を行った。このルーブリックでは、5章に記述したように従来のルーブリックとは異なる以下の新規性を有している。

- (1)重回帰分析により抽出した自由英作文の言語的特徴を特定要因の評価指標を評価の観点とする。
- (2)特定要因の評価指標を評価の観点とした分析的評価とする。
- (3)評価の観点に対する評価基準（到達度合い）を統計的に検証した定量的数値とする。
- (4)評価の観点に対する評価基準（到達度合い）の得点を自由英作文の言語的特徴を抽出した重回帰分析の各説明変数の標準化係数 β 値（全体的評価に対する評価の観点の貢献度）の比率により設定する。
- (5)評価を示す記述に加え、評価の観点の評価基準（到達度合い）ごとの定量的な数値を掲載し、学習者や指導者に提供することができる。

最終章では、各章の内容と要約し、その後、このようにしてデザインしたルーブリックが1章2節で示した研究の意義で示した事項を達成できるかどうかを確認することによって最終的な結論としたい。

1章では、本研究の対象となる学習者である初級および上級の日本人学習者をヨーロッパにおける外国語学習、教育および評価のガイドライン、CEFRの日本版であるCEFR-Jを基に規定し、前者をCEFR-J A2レベル、後者をCEFR-J B2およびC1レベルとした。このような学習によって記述される自由英作文の定義付けを行い、自由英作文というのは非常に複雑なメカニズムによって成立するライティング活動であることを示した。この自由英作文というのは、自身の考えや意見を発信する英語教育を求める社会的背景の中で、重要性が増している。しかしながら、その評価基準であるルーブリックが十分に整備されていないため、その教育の普及を阻害する1つの要因になっていることを述べた。そこで、本章において上述したような新規性を有する評価基準の指針であるルーブリックをデザインし、その有効性を検証することとした。このルーブリックによる研究の意義も示した。

2章では、全体的評価、分析的評価、特定要因の評価という3つの自由英作文の評価手法を示した。既存のルーブリックは、この全体的評価および分析的評価を基に作成されていることを示し、その長所と短所を分析した。本研究でデザインするルーブリックでは、特定要因の評価指標をルーブリックの評価の観点とする分析的評価を採用するという考え方を提示した。そのようなルーブリックにおいて、評価の観点となる特定要因の評価指標の選択の考え方を示すため、自由英作文の言語的特徴に関する4つの先行研究であるFerris(1994)、石田(2013)、杉浦(2008)、水本(2008)を分析した。その結果、自由英作文の言語的特徴を抽出するために重回帰分析を利用すること、決定係数 R^2 が最低限0.50

を超えることを目処に、説明変数候補（特定要因の評価指標）を選択することとした。これらの先行研究や自由英作文の言語的特徴以外の分析（例えば、特定要因の評価の1つである T-unit の有効性の検証）で使用された特定要因の評価指標を参考に、CEFR-J A2 レベル、CEFR-J B2 および C1 レベルの英語力を考慮して、その選択を行った。結果として、CEFR-J A2 レベルに関しては、Fluency（総語数、総文数）、文（1文の平均語数、Error-free T-unit 平均語数）、談話（結束数の合計、結束数の合計に占める指示 (R) および接続 (C) の割合）、語彙（GI、D、MTLD、P_Lex、S、平均語長）、リーダビリティ（FKG、ARI、CLI）の5つの視点から15指標を説明変数候補とした。CEFR-J B2 および C1 レベルに関しては、構成、内容、言語使用、Fluency（総語数、文数）、文（1文の平均語数）、語彙（GI、D、MTLD、P_Lex、S、平均語長）、リーダビリティ（FKG、ARI、CLI）の7つの視点から、15指標を説明変数候補とした。

3章では、CEFR-J A2 レベルの評価の観点を提案するために、まず、リーダビリティと語彙の豊かさ（多様性と広範さ）の視点から自由英作文を分析した。前者のリーダビリティの分析から、CEFR-J A2 レベルの学習者によって記述された自由英作文を理解するにはアメリカの小学校3年生から6年生程度の教育レベルが必要なことが判明し、リーダビリティを表わす特定要因の評価指標である CLI は自由英作文の言語的特徴を表わす可能性があることが示唆された。後者の語彙の豊かさの視点から分析では、語彙の広範さよりも多様性の方が、自由英作文の言語的特徴を表わす可能性が高いことが判明した。これらのリーダビリティと語彙の豊かさの2つの視点からの分析過程で、リーダビリティの評価指標である FKG と ARI、また、語彙の多様性を表わす評価指標である GI、語彙の広範さを表わす評価指標である D および S は、CEFR-J A2 レベルの自由英作文の言語的特徴の候補として妥当性を欠くことも判明した。そのため、2章で自由英作文の説明変数候補とした15の指標からこの5つの指標を除外した。残りの10指標を説明変数候補とし、自由英作文の全体的評価を目的変数とする重回帰分析を行った。結果として、総語数、CLI、MTLD、Error-free T-unit 平均語数の4指標が抽出された。この4指標に関して、ルーブリックの評価基準（到達度合い）の作成ため、統計的に有意になるように3段階の定量的な数値の算出を試みた。結果として、MTLD は有意な数値を得られず評価の観点として利用できなくなり、CEFR-J 1 A2 レベルの自由英作文のルーブリックでは総語数、CLI、Error-free T-unit 平均語数の3指標が評価の観点となることとなった。

4章では、CEFR-J B2 および C1 レベルの評価の観点を提案するために、まず、リーダビリティと語彙の豊かさ（多様性と広範さ）の視点から自由英作文を分析した。前者のリーダビリティの分析から、CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文を理解するにはアメリカの学校制度における10年生から11年生の教育レベルが必要なこと、また、リーダビリティの視点から分析する際に使用したリーダビリティの指標である ARI と CLI は自由英作文の言語的特徴を表わす特定要因の評価指標になりうる可能性があることが示唆された。語彙の豊かさ（多様性と広範さ）の視点から分析では、語彙の多

様性よりも広範さの方が CEFR-J B2 および C1 レベルの自由英作文の言語的特徴になりうる可能性が高いことが判明した。多様性と広範さの関係については、両者には重なる部分があり、中程度の相関関係があることが明らかとなった。これらのリーダビリティと語彙の豊かさの 2 つの視点からの分析過程で、リーダビリティの評価指標である FKG、また、語彙の多様性を表わす評価指標である GI、語彙の広範さを表わす評価指標である S は、CEFR-J B2 および C1 レベルの自由英作文の言語的特徴の候補として妥当性を欠くことも判明した。そのため、2 章で自由英作文の説明変数候補とした 15 の指標からこの 3 つの指標を除外した。残りの 12 指標を説明変数候補とし、自由英作文の全体的評価を目的変数とする重回帰分析を行った。結果として、Model 1 と Model 2 の 2 つのモデルが得られた。Model 1 では構成、内容、言語使用、Model 2 では言語使用と CLI が抽出された。ルーブリックの評価基準（到達度合い）の作成ため、統計的に有意になるように 3 段階の定量的な数値の算出を試みた。結果として、いずれの指標において統計的に有意になるように 3 段階の定量的な数値を得られた。従って、Model 1 を基にするルーブリックでは構成、内容、言語使用を、Model 2 も基にするルーブリックでは言語使用、CLI を評価の観点とすることとなった。

5 章では、Nishijima, et al(2007) の 7 項目に沿って、ルーブリックのデザインに関する考え方を示し、本章の冒頭に記述した新規性を取り入れたルーブリックをデザインした。CEFR-J A2 レベルのルーブリックでは以下のようになった。

表 5-4 CEFR-J A2 レベルの自由英作文のルーブリック（再掲）

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
情報性の豊かさ	情報量が十分に記述されており、内容がよくわかる。 得点：18 点（総語数 117 語以上）	伝えたい情報量がある程度記述されており、内容も一応理解できる。 得点：12 点（総語数 116 語～76 語）	伝えたい情報量が少なく、内容が理解しにくい。 得点：6 点（総語数 75 語以下）
語彙や文体の難易度	語彙や文体の難易度が高い。 得点：4.5 点（CLI:6.29 以上）	語彙や文体の難易度が普通である。 得点：3 点（CLI:6.28～5.53）	語彙や文体の難易度がやや低い。 得点：1.5 点（CLI:5.52 以下）
エラーが少なく、熟達した英文	エラーが少なく、十分に熟達した英文である。	ややエラーがあるものの、熟達した英文である。	エラーがあり、あまり熟達した英文になっていない。

得点：3点（EFTの平均 語数：8.35以上）	得点：2点（EFTの平均 語数：8.34～7.06）	得点：1点（EFTの平均 語数：7.05以下）
----------------------------	-------------------------------	----------------------------

同様に、CEFR-J B2 および C1 レベルのルーブリックを以下に示す。

表 5-9 Model 1 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック（再掲）

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
構成	序論で与えられたトピックを適切に把握し、本論に非常に適格につながる構成をしている。結論で本論と十分に首尾一貫したつながりを有している。 得点：3.0点 （構成の値：3.90以上）	序論で与えられたトピックを把握し、本論につながる構成をしている。結論で本論と首尾一貫したつながりを有している。 得点：2.0点 （構成の値：3.89～3.43）	序論で与えられたトピックを把握し、本論につながる構成を成していない。結論で本論と首尾一貫したつながりを有していない。 得点：1.0点 （構成の値：3.42以下）
内容	本論の主題部分の内容が非常に適切であり、その支持部分の内容も極めて適切である。 得点：7.5点 （内容の値：4.90以上）	本論の主題部分の内容が適切であり、その支持部分の内容も適切である。 得点：5.0点 （内容の値：4.89～3.48）	本論の主題部分の内容が必ずしも適切ではなく、その支持部分も適切ではない。 得点：2.5点 （内容の値：3.47以下）
言語使用	語彙や文法の誤りがほとんどみられない。 得点：7.5点 （言語使用の値：4.97以上）	語彙や文法の誤りがややみられる。 得点：5.0点 （言語使用の値：4.96～3.73）	語彙や文法の誤りがかなりみられる。 得点：2.5点 （言語使用の値：3.72以下）

表 5-10 Model 2 の CEFR-J の B2 および C1 レベルの自由英作文のルーブリック（再掲）

評価の観点	評価基準（到達の度合い）		
	上位者	中位者	下位者
言語使用	語彙や文法の誤りがほとんどみられない。 得点：6点 （言語使用の値：4.97以上）	語彙や文法の誤りがややみられる。 得点：4点 （*言語使用の値：4.96～3.73）	語彙や文法の誤りがかなりみられる。 得点：2点 （*言語使用の値：3.72以下）
語彙や文体の難易度	語彙や文体の難易度が高い。 得点：3点 （CLI:12.24以上）	語彙や文体の難易度が普通である。 得点：2点 （CLI:12.23～10.83）	語彙や文体の難易度がやや低い。 得点：1点 （CLI:10.82以下）

6章では、提案したルーブリックの使い方とその有効性を検証した。CEFR-J A2 レベルのルーブリック、CEFR-J B2 および C1 レベルの2つのルーブリックによって算出される得点と自由英作文の全体的評価とは、いずれも正の高い相関関係が認められ、その有効性が検証された。

このようにしてデザインしたルーブリックが1章2節で示した研究の意義を達成できるかどうかを以下に検証する。1章2節で提示した CEFR-J A2 レベルのルーブリックに関する研究の意義を以下に示す。

- ① ‘Independent WRITING Rubrics’ では一律に評価の低い（1点）と判定された自由英作文の評価の峻別が可能となる。
- ②自由英作文の言語的特徴に基づく評価の観点は学習者に学習の指針を提供することになる。
- ③指導者による自由英作文の評価の際の効率性を高めることができる。
- ④自由英作文の言語的特徴に基づく評価の観点の定量的な数値情報は指導者に指導の指針を提供することになる。

表 5-4 の CEFR-J A2 レベルのルーブリックを使用すると、「① ‘Independent WRITING Rubrics’ では一律に評価の低い（1点）と判定された自由英作文の評価の峻別が可能となる」に関しては、3つの評価の観点から、最低点 8.5 点から最高点 25.5 点まで採点が可能となり、「Independent WRITING Rubrics’ では一律に評価の低いと評価される自由英作文の評価の峻別は十分可能である」と考える。「②自由英作文の言語的特徴に基づく評価の観点は学習者に学習の指針を提供することになる」に関しては、「情報性の豊かさ」

「語彙や文体の難易度」、「エラーが少なく、熟達した英文」の上位者の条件を満たせば、より質の高い自由英作文となることが学習者にも容易に理解することができる。さらに、「情報性の豊かさ」では、総語数を増やすことが大切であることを学習者に指示してもよい。「語彙や文体の難易度」は CLI というリーダビリティの指標で、その自由英作文を理解するには、当該数値のアメリカの学年の教育レベルが必要であるという意味合いを提供するとともに、このことは、英文を書く際に、読み手を意識して英文を書くことが重要であることを表しているという指針を学習者に提供してもよい。「エラーが少なく、熟達した英文」とは Error-free T-unit の平均語数を向上させることで、エラーを減らし、複文構造の英文を使用することという指針を学習者に提供することができる。「③指導者による自由英作文の評価の際の効率性を高めることができる」に関しては、評価者は評価の観点の数値を算出しさえすれば、その数値を基に、対象とする自由英作文を評価することが可能となるため、評価の効率性は高まるはずである。「④自由英作文の言語的特徴に基づく評価の観点の定量的な数値情報は指導者に指導の指針を提供することになる」に関しては、各評価の観点の定量的な数値の向上をはかるような指導を工夫することによって達成されると考える。例えば、「情報性の豊かさ」である総語数を増やすために、アイディアの出し方の指導方法の工夫などを考えることができるはずである。

次に CEFR-JB2 および C1 レベルのルーブリックに関する研究の意義を以下に示す。

- ① ‘Independent WRITING Rubrics’ では高評価として 4 点ないし 5 点と判定された自由英作文をさらに峻別できる評価が可能となる。
- ②自由英作文の言語的特徴に基づく評価の観点は学習者に学習の指針を提供することになる。
- ③指導者による自由英作文の評価の際の効率性を高めることができる。
- ④自由英作文の言語的特徴に基づく評価の観点の定量的な数値情報は指導者に指導の指針を提供することになる。

表 5-9 Model1 の CEFR-JB2 および C1 レベルのルーブリックを使用すると、「① ‘Independent WRITING Rubrics’ では高評価として 4 点ないし 5 点と判定された自由英作文をさらに峻別できる評価が可能となる」に関して、4 つの評価の観点から、最低点 6.0 点から最高点 18.0 点まで採点が可能となり、「Independent WRITING Rubrics’ では一律に高評価として 4 点ないし 5 点と判定される自由英作文の評価の峻別は十分可能である」と考える。「②自由英作文の言語的特徴に基づく評価の観点は学習者に学習の指針を提供することになる」に関しては、「構成」「内容」、「言語使用」の上位者の条件を満たせば、より質の高い自由英作文となることが学習者にも理解することができる。さらに、この 3 つの評価の観点を算出するために 6 章に提示した評価シート（表 6-3～表 6-8）も学習者に活用させることによっても、自由英作文に対する学習の指針の提供は促進されるはずである。「③指導者による自由英作文の評価の際の効率性を高めることができる」に関しては、評価者は評価の観点の数値を算出しさえすれば、その数値を基に、対象とする自由英作文

を評価することが可能となるため、評価の効率性は高まるはずである。「④自由英作文の言語的特徴に基づく評価の観点の定量的な数値情報は指導者に指導の指針を提供することになる」に関しては、各評価の観点の定量的な数値の向上をはかるような指導を工夫することによって達成されたと考える。例えば、「構成」であれば、与えられたトピックを踏まえ、うまく本論につながるような序論を書かせる指導が大切であるという指針を提供することができるはずである。

表 5-10 Model 2 の CEFR-JB2 および C1 レベルのルーブリックを使用すると、「① ‘Independent WRITING Rubrics’ では高評価として 4 点ないし 5 点と判定された自由英作文をさらに峻別できる評価が可能となる」に関して、2 つの評価の観点から、最低点 3.0 点から最高点 9.0 点まで採点が可能となり、「Independent WRITING Rubrics’ では一律に高評価として 4 点ないし 5 点と判定される自由英作文の評価の峻別は十分可能である」と考える。「②自由英作文の言語的特徴に基づく評価の観点は学習者に学習の指針を提供することになる」に関しては、「言語使用」、「語彙や文体の難易度」の上位者の条件を満たせば、より質の高い自由英作文となることが学習者にも理解することができると思う。さらに、「言語使用」の評価の観点を算出するために 6 章に提示した評価シート（表 6-9）も学習者に活用させることによって、自由英作文に対する学習の指針の提供は促進されるはずである。「語彙や文体の難易度」は CLI というリーダビリティの指標で、その自由英作文を理解するには、当該数値のアメリカの学年の教育レベルが必要であるという意味合いを提供し、上級の学習者であるため、CLI が上位者の基準値 12.24 を越える英文を読解させることによって、語彙や文体を学ばせることもできる。「③指導者による自由英作文の評価の際の効率性を高めることができる」に関しては、評価者は評価の観点の数値を算出しさえすれば、その数値を基に、対象とする自由英作文を評価することが可能となるため、評価の効率性は高まるはずである。「④自由英作文の言語的特徴に基づく評価の観点の定量的な数値情報は指導者に指導の指針を提供することになる」に関しては、各評価の観点の定量的な数値の向上をはかるような指導を工夫することによって達成されたと考える。例えば、「CLI」であれば、上述したように上位者の基準値 12.24 を越えるような英文を学習者に読解させ、語彙や文体の難易度を上げる指導を工夫することができる。

このように本研究で提案した自由英作文の 3 つのルーブリックはその研究の意義を達成したと考える。しかしながら、今後、このルーブリックを普及させる際には、当初は様々な疑問が投げかけられるはずである。5 章 2 節に提示した A 先生と同様の批判を受けるのは至極当然である。そのような批判には、「自由英作文をどのように指導していますか」という問いかけをし、まずは、その回答をいただく。その回答がどのようなものであろうとも、このルーブリックを活用すれば、回答していただいた指導方法を補完することになり、おそらく、A 先生のように新たな指導法を次々と考案、工夫するトリガーとなるはずである。そのようなトリガーが頻発すれば、筆者としては望外の喜びであり、提案したルーブリックが、停滞する日本における CEFR-J A2 レベルと CEFR-J B2 および C1 レベルの学習者

に対する自由英作文の教育を革新するツールとなりうることを強調して結論とする。

7.2 今後の課題

本研究では、CEFR-J A2 レベルと CEFR-J B2 および C1 レベルの学習者に対する自由英作文の評価基準であるルーブリックをデザインすることを主目的としたため、その評価を活用した実際の指導方法には、一定の示唆を提供するに留まった感がある。今後は、これらのルーブリックの評価の観点、評価基準（到達の度合い）を活用した指導方法の開発を探求していくことが重要であると考えられる。また、3章2節で記述した CEFR-J A2 レベルの語彙の豊かさの分析や4章2節で記述した CEFR-J B2 および C1 レベルの語彙の豊かさの分析、さらに小島（2013）の分析を考慮すると、CEFR-J B1 レベル（TOEIC 550 点～785 点）以下の CEFR-J A2 レベルの自由英作文では語彙の多様性が、その全体的評価に影響を与え、CEFR-J B1 レベルを越えると語彙の広範さがその全体的評価に影響を与えるという知見が得られている。このような語彙の分析から得られた知見も提案したルーブリックによる指導方法の開発に活用していくこととしたい。

本研究は日本人という完全に EFL の環境で英語を学んだ学習者を対象として、ルーブリックの提案を行ったが、その枠組みを広げ、より ESL に近い環境で英語を学んだヨーロッパの人達、例えば、オランダ人やドイツ人のように高度な英語力を有すると言われる英語の非母語話者では、提案した上級者向けのルーブリックが適用できるかどうかの検討も考えてみたい。石田（2013）は英語の非母語話者（日本人大学生および大学院生）と母語話者によって記述された自由英作文を比較し、総語数、異なり語数、GI と 2 語表現の総数が両者を判別する言語的特徴であるとしている。英語の母語話者並の運用能力を有するヨーロッパの人達と日本人の上級英語学習者によって記述された自由英作文においても、石田（2013）のような差異が存在するかどうか、現状では明らかとなっていない。その分析を行う中で、この日本人上級者学習者向けのルーブリックの更なる拡大版の提案を今後の研究対象ともしていきたいと考える。

資料 1-1 CEFR-J A2 レベルの学習者によって記述された自由英作文の事例

(1) ライティング力下位者による自由英作文の事例

I will write about social problem in Japan. I think “Poi sute” is a serious problem. “Poi sute” is not to desert trashes to trash box. I have three ideas for this problem. First, it is too bad to desert trashes around the city, because it makes dirty the city. Second, I want to clean the cities around the world. If it become clean, people will be happy. Third, I think that it is common sence to desert trashes to trash box.

In summary, I think if you find the man who deserting trashes, you should warn him. Same women.

(2) ライティング力中位者による自由英作文の事例

Japan has a lot of gabages. Of course, the world has these too. But Japan is a small country. So we don’ t have places to desert gabages. But Japanese don’ t think about it. My hometown is near the sea. The sea isn’ t clear. Because, people don’ t take off their gabages. When I saw the scene, I was sad. Why people don’ t clean the sea? I thought we must take action and we must keep to protect beautiful places. But in these days, people began to think about many gabages. Some people cleaned may places. Other people didn’ t desert a lot of gabages. I think it is good all Japanese think about it.

(3) ライティング力上位者の自由英作文の事例

I think that Japanese has many problems. For example Japan has many trash, I think this problem is very serious. I see many duty trash everyday. The trash boxies which is in station always are full of trash. Many cans and papers and so on in it. I sometimes see threw trash. I don’ t like the sction and I think that not to do that. I think threw trash is cause for the part of problem. If threw trash were stop, Japan would be beautiful country. However, many people do that. I feel sad. By the way, how many people do they know “NO Rezibukuro” . I know that. When I buy something, I tell “no rezibukuro” . This is good for the earth. I think that Japanese trash problem is world problem.

資料 1-2 CEFR-J B2 および C1 レベルの学習者によって記述された自由英作文の事例

(1) CEFR-J B2 レベルの学習者によって記述された自由英作文の事例

I disagree with the idea that people should follow the customs of another country when they move to the new country. Even if people move out of their homeland, their food, clothes and houses are still symbols of their native culture. People's customs connected directly with their living necessities have been created with their traditional culture.

First, here is a good example of food. Chinese people have been moving to all over the world. Making their local communities like China-towns abroad, they not only keep their custom of eating Chinese food, but also serve the cuisine to the local citizens. They prefer to retain their native tastes for foods rather than changing into those of the new country.

Second, let me talk about clothes. People of a Himalayan kingdom, Bhutan, wear their traditional clothes wherever they go even out of the country. They are strongly proud of their native costumes because they believe that their traditional clothes represent the spirit of the country.

Finally, both Chinese and Bhutanese people are mostly Buddhist. They do not decorate Christmas ornaments on their houses even in a western country while the local residents do so as their customs in December.

I disagree with the opinion that people should follow the customs of another country when they move to the new country. As the examples of the Asian people shown above, keeping their native customs abroad is much more important than changing their customs. They are strongly proud of their native customs wherever they live even abroad.

(2) CEFR-J C1 レベルの学習者によって記述された自由英作文の事例

Some people say that when you move to another country, you should follow the customs of the new country. Other people say that you need not to follow their customs. Personally I agree with the idea that you should follow the customs of the new country for the following three reasons.

Firstly, it will help you to construct a good relationship with people you meet in the new country. The customs are deeply connected to the people's way of thinking. Following the customs demonstrates that you understand and respect people in the country, which will make a good impression on them. A good impression will lead to a friendly and comfortable relationship.

Second, it will smoothen the process of assimilating yourself to the new country. The customs are inextricably linked with the lifestyles and culture of the country.

Following the customs will naturally lead to the understandings of the new lifestyles and culture, facilitating your adaptation to the new country.

Third, it will improve your communication skills with people you meet in the new country. The customs constitutes the background of the language of the country. Following the customs will deepen your understandings of the language, which will increase your fluency in it. The better language skills will enable you to have a wider and deeper communication with new neighbors and friends.

In conclusion, I believe that when you move to another country, you should follow their customs for the three reasons: a rapport with new people, better adaptation to the new country, and better communication with new people.

資料 1-3 2013 年自由英作文出題大学

大学名	課題
愛知教育大学 (前期日程／教育, 英語専修・英語専攻・国際文化コース志願者)	あなたにとって最も思いで深い場所はどこか。その理由も明記しながら、100 語程度の英語で書け。
愛知県立大学 (前期日程／外国語・日本文化・教育福祉・看護)	Write a short essay in English (100-150 words) on what you hope to gain from your time at university.
旭川医科大学 (前期日程／医学部)	Do you agree with the following statements? University students aged 20 and should be allowed to drink alcohol on campus during school festivals. Write an essay in English. If you agree, explain the advantages. If you disagree, explain the disadvantages. Use reasons and examples when giving your opinion.
岩手大学 (前期日程／人文社会科学)	英語の授業で、「東北地方の高等学校は冷房使用を禁止すべきである」というテーマのディベートに挑戦し、あなたは賛成派の立場になったとする。(1)あなたが賛成の立場をとる理由、(2)あなたが(1)の立場で述べた賛成の立場に対し、予想される反論、(3)予想される反論に対するあなたの再反論、という構成でスピーチ原稿を準備せよ。以下のキーワードを適宜参考にしながら、(1)～(3)ともに 40 語(words)以上の英語でまとめること (キーワードは使用しても使用しなくてもどちらでもいい)。
大阪大学 (後期日程／外国語学部)	次の日本語を読んで、設問に答えなさい。 (1)このエッセイに書かれている「私」派と「彼女」派の 2 つのタイプについて 100 語程度の英語で説明しなさい。 (2)現代社会における「私」派と「彼女」派のメリットとデメリットについて 100 語程度の英語で書きなさい。
大阪教育大学 (前期日程／教育・学校英語)	To cope with the problem mentioned by the author, what solution would you suggest? Explain it in 80 to 100 English words.
小樽商科大学 (前期日程／商)	Write two short essays in English to answer both part (a) and part (b) of this question.

	<p>(a) Imagine that the three birds in the picture below could talk. Write a story explaining who they are, why they met and what they are talking about. (About 60 words)</p> <p>(b) The following information in Table 1 shows data about leisure patterns in Japan by age group. Write an essay in English describing these patterns. (About 60 words)</p>
<p>香川大学 (前期日程／教育・法・経済・医)</p>	<p>次の英文の指示に従って、自分の考えを 12 行程度の英文でまとめなさい。</p> <p>Write about an experience you had in which you tried something new. Explain what you did and what you learned from this experience.</p>
<p>鹿児島大学 (前期日程／法文・教育・理・医・歯)</p>	<p>Write your answer in English to the following question in 80 to 100 words. What skills do you think are needed to succeed in society? Give two most needful skills and your reasons for choosing them.</p>
<p>北九州大学 (前期日程／外国語)</p>	<p>次の英文の質問に、60 語以上 80 語以下の英文で答えなさい。</p> <p>If you could publish a book on any topic, what would you write about? Give your book a title, and briefly describe what you would write about.</p>
<p>九州大学 (前期日程／文・教育・法・経済・理学・医・歯・薬学・工学・芸術工学・農学)</p>	<p>次の英文の指示に従い、100～120 語程度の英語の文章を書きなさい。</p> <p>It is sometimes important to be able to work with a group of people. Provide an example from your own experiences of working with others and discuss the benefits and difficulties of such group work.</p>
<p>京都工芸繊維大学 (前期日程／工芸科学)</p>	<p>Answer the following questions in English in 150-200 words.</p> <p>Questions: What do you think about the generation gap between young people and the older generation in Japan? Explain your opinions providing concrete examples and reasons.</p>
<p>京都府立医科大学 (前期日程／医)</p>	<p>Answer the following question in about 150 English words.</p> <p>The University of Tokyo announced last year that they would change the start of the academic year from April to September. Since then, there have been heated discussion around the country, which means that there are both positive and negative</p>

	<p>responses to the new system. What do you think of this? First of all, state if you are “For” or “Against” this change, and then explain logically your opinion about this change in the university calendar. Support your view with two or three specific examples.</p>
<p>高知大学 (前期日程／教育)</p>	<p>下記の意見に対して、賛成か反対か、あなたの立場を明らかにし、その理由とともに 100 語以内の英語で書きなさい。 “Top high school students should be allowed to enter university a year early.”</p>
<p>神戸市立外国語大学 (前期日程／外国語)</p>	<p>Write approximately 100 words in English in answer to the following question: There is a saying, “We all make mistakes” and this might be good for us. What are the benefits of making mistakes?</p>
<p>公立はこだて未来大学 (前期日程／システム情報科学)</p>	<p>Opinions What are three skills that a good scientist needs in the 21st century? Write one paragraph (at least 60 words) in English explaining why these skills are important. Explanatory Skills Write one paragraph (at least 60 words) in English to convince the author of [1] that technology is a good thing. Including at least three reasons. [1]:問題番号 1 Personal Response Think about your own life and write one paragraph (at least 60 words) in English describing how technology has caused problems for you or someone you know personally .</p>
<p>国際教養大学 (A 日程／国際教養)</p>	<p>Based on the article*, write an essay of 250 words or more explaining your beliefs and experiences on the topic of people’s need for sleep. *Rethinking Sleep という The New York Times の記事</p>
<p>滋賀医科大学 (前期日程／医)</p>	<p>Read the following Japanese passage on laughter carefully, and, with this in mind, write an English essay of about 150 words, where you analyze the possible causes or circumstances that might produce laughter and the nature of that laughter, illustrating with your own experiences and thoughts about it in detail.</p>

静岡大学 (前期日程／人文・教育・情報)	Imagine you are the principal of your high school, and you have the power to change anything you want. What improvements would you make? Include the details about the changes and the reasons for them. Write about 120 words in English.
島根大学 (前期日程／法文・教育・総合理工・生物資源科学)	Some people like to travel with a friend. Other people prefer to travel alone. Which do you prefer? Explain the reasons in three to five English sentences.
電気通信大学 (前期日程／情報理工)	次の二つの質問から一つ選んで、二つ以上の理由を具体的に挙げて英語で答えなさい。 1. Do you prefer traveling in a group or traveling alone? Give your reasons. 2. In your opinion, what should you be done to encourage people to save electricity?
東京大学 (前期日程／文科一～三類・理科一～三類)	これまで学校や学校以外の場で学んできたことなかで、あなたが最も大切だと思うことは何か、またそれはなぜか。50～60語の英語で答えよ。ただし、英語に関すること以外について述べること。
東京大学 (後期日程／文科一～三類・理科一～三類)	Suppose native speakers of English used Japanese words and phrases in the same way Japanese use wasei eigo. Explain how you would react, and why, in 80 to 120 English words. As much as you can, avoid copying from the given text.
東京農工大学 (前期日程／農・工)	In modern society, news is available from a number of media. Which one do you like best? Give two reasons why you prefer this particular medium to the others. Write one or more sentences in English (30-40 words).
新潟大学 (前期日程／人文・教育・法・経済・理・医・歯・工・農)	次の質問に 100 語 (100 words) 程度の英文で答えなさい。 If you could change one thing about yourself, what would it be and why?
浜松医科大学 (前期日程／医)	Write an essay in English on the following topic. Your essay should be: 1.be a minimum of 150 words, 2.be written using paragraph form, 3.have a minimum of three paragraphs,

	<p>4.have a clear introduction, body and conclusion, 5.leave a one-line space between each paragraph.</p> <p>Do not double-space your essay; write on every line. In your essay, your ideas should be clearly expressed.</p> <p>The Hippocratic Oath is the traditional pledge made by doctors upon graduating from their medical studies. It describes what doctors see as their duties and obligations with regard to their patients and the medical profession as a whole. Describe what you think are a doctor’s main duties(apart from “helping people” or “saving lives”), obligations and responsibilities. Think outside the box, that is, be creative in formulating your answer.</p>
<p>一橋大学 (前期日程／商・経 済・法・社会)</p>	<p>Write 120 to 150 words of English on one of the statements below. Indicate the number of the statement you have chosen. Also, indicate the number of words you have written at the end of the composition.</p> <p>1 Government must protect and promote so-called “endangered” languages that are disappearing quickly today. Explain why this is true.</p> <p>2 Abraham Lincoln once said, “It has been my experience that people who have no vices have very few virtues.” Explain why this is true.</p> <p>3 Recently, many governments around the world are legalizing marriage between two people of the same sex. Legalizing same-sex marriage is a good idea in Japan. Explain why this is true.</p>
<p>一橋大学 (後期日程／経済)</p>	<p>Write 120 to 150 words of English about one of the topic below. Indicate the number of the statement you have chosen. Also, indicate the number of words you have written at the end of the composition.</p> <p>1 Herry Ford once said, “A business that makes nothing but money is a poor business.” Explain why this is true.</p> <p>2 Some people believe that there are no minorities living in Japan. Explain why this statement is false.</p> <p>3 Let’s imagine that we have discovered life on Mars that is</p>

	more intelligent than human beings. Explain why this could be dangerous to us.
兵庫県立大学 (前期日程／経済)	Write an essay in English about the following question. (about 100 words) Which do you think you would enjoy the most: traveling alone, travel with one or just a few friends, or traveling with a large group of friends? Why?
弘前大学 (前期日程／人文・医・教育)	If you could have dinner with a famous person, past or present, who would it be? Why? Answer in about 70 English words.
広島大学 (前期日程／文・教育・法・経済・理・医・歯・工・生物生産)	あなたが記録映画（ドキュメンタリー）を制作するとしたら、何について制作しますか。その理由や重要性を含めて 90 語程度の英語で書きなさい。
福岡教育大学 (前期日程／教育)	初等教育教員養成課程英語選修志願者 Imagine that you will visit a high school in an English speaking country and that you are going to introduce some aspects of Japanese culture. What aspects of Japanese culture would you like to introduce? Decide on two aspects. In English, describe these two aspects and your reasons, using approximately 150 words. 中等教育教員養成課程英語専攻志願者 In most high schools in Japan, only one foreign language is taught, English. If an additional foreign language could be taught in your school, what language would you like to be taught? In English, describe the language you would like to be taught and why, using approximately 150 words.
福島大学 (前期日程／人文社会科学群)	次の 2 つの英語の質問から 1 つ選び、選択した質問の番号を記入のうえ、100 語程度の英語で自分の考えを書きなさい。 (1) What is the most difficult problem that high school students face these days? (2) Which do you think is more important, money or friendship?
北海道大学 (前期日程／文・教	People rescued in the mountains in Japan should pay the full costs of their rescue. Do you agree or disagree? Explain why in

育・法・経済・医・歯・ 獣医・水産)	70-100 words giving two reasons to support your opinion.
宮城教育大学 (前期日程／教育)	以下の三つのトピックから一つ選んで、自分の考えを 5 行程度の英文で述べなさい。 1) Recently, it is said that many people sitting in the priority seats on trains and buses won't give up their seat for older adults. What do you think about that? 2) There is a famous proverb in English: An apple a day keeps the doctor away. What do you do to "keep the doctor away" 3) Some people say we should spend more money to explore the universe. What do you think about this?
宮崎大学 (前期日程／教育文化・農)	以下の英文を読んで、その指示に従って 10 行程度の文章を英語で書きなさい。 What is one of your favorite subjects at school? Describe the subject and explain why you particularly like it?
琉球大学 (前期日程／法文・観光産業科学・医・教育)	次の問に答えなさい。 あなたは、高校 3 年生です。夏休みが終わったあとで、友人から高校をやめたいと相談されました。もしあなたが、その友人に高校は卒業してほうがいと説得するとしたら、どのように説得しますか。150 字程度の英語で述べなさい。なお、内容も含め、文法やスペリングも採点の対象となります。

業績一覧

投稿論文

- 平林健治. (2014). 「CEFR-J A2 レベルの自由英作文のリーダビリティによる分析」. 『言語文化学会論集』 42, 131-145.
- 平林健治. (2015 a). 「CEFR-J A2 レベルの自由英作文の言語的特徴の分析」『関東甲信越英語教育学会 学会誌 KATE Journal』, 29 号, 1 - 14.
- 平林健治. (2015 b). 「CEFR-J A2 レベルの自由英作文の語彙の豊かさに関する研究」『中部地区英語教育学会紀要』 第 44 号, 17 - 24.
- 平林健治. (2015 c) 「日本人上級英語学習者の自由英作文のリーダビリティによる分析：パイロットスタディ」『言語文化学会論集』 44, 133-142.
- 平林健治. (2016 a). 「日本人上級英語学習者の自由英作文の語彙の豊かさに関する研究」『関東甲信越英語教育学会 学会誌 KATE Journal』, 30 号, 29-42.
- 平林健治. (2016 b). 「日本人上級英語学習者の自由英作文の言語的特徴の分析」『中部地区英語教育学会紀要』 第 45 号, 23-30.
- 平林健治. (2016 c) 「CEFR-J A2 レベルの自由英作文のルーブリックに関する研究」『日本教科教育学会誌』 第 39 巻 2 号, (掲載決定)

学会発表

- 平林健治 「CEFR-J A2 レベルの自由英作文の語彙の豊かさに関する研究」, 第 44 回中部地区英語教育学会山梨大会 2014 年 6 月
- 平林健治 「日本人上級英語学習者の自由英作文の言語的特徴の分析」, 第 45 回中部地区英語教育学会和歌山大会 2015 年 6 月
- 平林健治 「日本人上級英語学習者の自由英作文の語彙の豊かさに関する研究」, 関東甲信越英語教育学会第 39 回山梨研究大会 2015 年 8 月

謝辞

本論文の作成にあたり、多くの方々の御指導と御協力をいただきました。心より感謝申し上げます。

まず、指導教員の前野隆司先生には大変、お世話になりました。遠方で勤務している関係上、あまり日吉に来ることができない環境にいる私にも、様々な形で刺激を与えていただき、時に、厳しいことばを頂き、それに見合うだけの研究にしていこうという意欲を掻き立てていただきました。今回の研究につきましても先生の鋭い指摘によって、その方向性が定まりました。先生のあの指摘がなければ、この博士論文を書き上げることはできなかったと思います。改めて、お礼申し上げます。

副指導教員の高野先生からも、下読み会で大変貴重な意見を頂き、感謝致します。私の勘違いで、資料が不十分なまま下読み会に臨んでしまい、大変御迷惑をおかけしました。それにもかかわらず、論文を評価していただくおことばを頂いた時、この研究に取り組んでよかったことを痛感しました。有難うございました。

また、白坂先生にも副指導教員として、お世話になりました。白坂先生との下読み会では、資料をある程度、整えて臨むことはできましたが、自分の研究を人にどのように分かりやすく、かつ、正確に伝えるかという自分の最も苦手な部分の向上につながる大変貴重なアドバイスを頂き、感謝致します。

静岡文化芸術大学の横田秀樹先生とは、15年以上の研究での付き合いがありますが、今回、私の勝手を聞き入れて下さり、副指導教員を務めていただきました。英語教育の学会でも、よくお話をさせていただいておりますし、紙面で感謝を記述するのは照れてしまいますが、先生からいただく、優しいことばでの鋭い指摘はとても刺激的です。今回の研究でも、いつもと変わらぬ指摘をいただき、有難うございました。

本研究では、アクエアリーズという英語学校の植田一三先生と、同校の講師の方、また、同校の受講生の方々にも、協力いただきました。本当に有難うございました。

SDM に入学し、既に4年が経過しましたが、あまり、日吉に行くことができなかったのが後悔されます。ただ、回数は少なくとも、そこで学ぶ若くて優秀な院生の方々を目にすると自分もまだまだ、頑張ろうという意欲がわきました。SDM 入学し、本当によかったです。深く感謝致します。

参考文献

- 相澤一美. (2001). 「ライティング能力の測定と評価」小室俊明 (編)『英語ライティング論』(pp. 62-72). 東京: 河源社.
- 相澤一美・石川慎一郎・村田 年・磯 達夫・上村俊彦・小川貴宏・清水伸一・杉森直樹・羽井佐昭彦・望月正道. (2005). 『大学英語教育学会基本語リストに基づく JACET 8000』. 東京: 桐原書店.
- 青谷正安. (2012). 『英語学習論—スピーキングと総合力』. 東京: 朝倉書店.
- Arapoff, N. (1967). "Writing: A Thinking Process," *TESOL Quarterly*, 1, 2, Jun., 33-39; EFT, 13, 3 & 4, special Issue (Part 2), (1975), 233-237.
- Astika, G. (1993). Analytical assessments of foreign students writing. *RELC Journal*, 24(1), 61-72.
- 馬場千秋. (2008) 「日本人英語学習者のライティング能力—英作文と日本語作文の関係」日英言語文化研究会 (編)『日英の言語・文化・教育—多様な視座を求めて』(pp. 307-317). 東京: 三修社.
- 馬場千秋. (2010). 「ライティング指導で求められているもの」木村博是・木村友保・氏木道人 (編)『英語教育学体系 第10巻 リーディングとライティングの理論の実践』(pp. 109-134). 東京: 大修館書店.
- Bardovi-Harlig, K. & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, 11, 17-34.
- Connor, U. (1990). Linguistic / rhetorical measures for international student persuasive writing. *Research in the Teaching of English*, 24, 67-87.
- Cumming, A. (1989). Writing expertise and second-language proficiency. *Language Learning*, 39, 81-141.
- Daller, H., & D, Phelan. (2007). What is in a teacher's mind? Teacher ratings of EFL essays and different aspects of lexical richness. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 93-115). Cambridge University Press.
- Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4(2), 139-155.
- Fathman, A. K. & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press.
- Ferris, D. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly*, 28, 414-420.
- Flower, L. & Hayes. (1981). A cognitive process theory of writing. *College Composition*

- and Communication*, 32(4), 365-387.
- Gaies, S. J. (1980). T-unit analysis in second language research: Applications problems and limitations. *TESOL Quarterly* 14, 53-60.
- Gosten, H. (1996). Verbal reports of Japanese novices' research writing practices in English. *Journal of Second Language Writing*, 5(2). 109-128.
- 河原俊昭. (2013). 「英語教育の到達目標としての TOEIC と CEFR の有効性の考察」. 京都光華女子大学・英語英米文学会 (編). 『国際化時代の言語文化』 (pp. 21-72). 大阪: 大阪教育図書.
- 萩野俊哉. (1998). 『ライティングのための英文法』. 東京: 大修館書店.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. New York: Longman. (安藤貞雄・多田保行・永田龍男・中川憲・高口圭輔訳. (1997). 『テキストはどのように構成されるか— 言語の結束性 —』. 東京: ひつじ書房.
- Halliday, M. A. K., & Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Deakin University Press. (笈壽雄訳. (1991). 『機能文法のすすめ』. 東京: 大修館書店.)
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar Second Language*. Edward Arnold Limited. (山口登・笈壽雄訳. (2001). 『機能文法概説—ハリデー理論への誘い—』. 東京: くろしお出版.)
- Hamamoka, Y. (1999). Can Japanese students pass the entrance examinations for Universities? *Journal of Pan-Pacific Association of Applied Linguistics*, 3, 273-279.
- Hamp-Lyons, L. (1991 a). Scoring procedures for ESL contexts. In Hamp-Lyons, L., editor, *Assessing second language writing in academic contexts*. Norwood: Ablex, 241-276.
- Hamp-Lyons, L. (1991 b). Basic Concepts. In Hamp-Lyons, L., editor, *Assessing second language writing in academic contexts*. Norwood: Ablex, 5-15.
- 橋内武. (1995). 『パラグラフ・ライティング入門』 東京: 研究社.
- 葉田野不二美・馬場千秋. (2006). 「大学生の英作文に見られる語彙に関する研究: 語彙サイズ別語彙の使用度」『外国語教育学会紀要』 No. 9, 56-71.
- 樋口晶彦. (2010). 「パラグラフ・ライティングに基づいたプロセス・ライティング」. 木村博是・木村友保・氏木道人 (編). 『英語教育体系 第10巻 リーディングとライティングの理論と実践—英語を主体的に「読む」・「書く」』 (pp. 198 - 213). 東京: 大修館書店.
- 廣森知人. (2015). 『英語学習のメカニズム—第二言語習得研究にもとづく効果的な勉強法』 東京: 大修館書店.
- 平林健治. (2003). 「自由英作文の分析: 総合的評価に対する分析的評価の変数と結束性の

- 視点から」『中部地区英語教育学会紀要』第 33 号, 169-176.
- 平林健治. (2004). 「FW 評価における T-unit の妥当性」『関西教育学会紀要』第 28 号. 91-95.
- 平林健治. (2005). 「日本人初級英語学習者の Free Writing 能力の要因に関するモデルの構築」. 名古屋大学大学院国際開発研究科提出博士論文.
- 平林健治. (2014). 「CEFR-J A2 レベルの自由英作文のリーダビリティによる分析」. 『言語文化学会論集』 42, 131-145.
- 平林健治. (2015 a). 「CEFR-J A2 レベルの自由英作文の言語的特徴の分析」『関東甲信越英語教育学会 学会誌 KATE Journal』, 29 号, 1 - 14.
- 平林健治. (2015 b). 「CEFR-J A2 レベルの自由英作文の語彙の豊かさに関する研究」『中部地区英語教育学会紀要』第 44 号, 17 - 24.
- 平林健治. (2015 c) 「日本人上級英語学習者の自由英作文のリーダビリティによる分析：パイロットスタディ」『言語文化学会論集』 44, 133-142.
- 平井明代(編). (2012). 『教育・心理系研究のためのデータ分析入門—理論と実践から学ぶ SPSS 活用法』 東京：東京図書.
- 広瀬恵子. (1999). 「L2 ライティング研究の概観と英語のライティング指導への示唆」『愛知県立大学外国語学部紀要（言語・文学編）』第 31 号, 41-65.
- Homburg, T. (1984). Holistic evaluation of ESL composition: Can it be validated objectively? *TESOL Quarterly*, 18, 87-106.
- Hulstijn, J. H., Schoonen, R., De Jong, N. H., Steinel, M. P., & Florijn, A. (2012). Linguistic competences of learners of Dutch as a second language at the B1 and B2 levels of speaking proficiency of the Common European Framework of Reference for Languages (CEFR). *Language Testing*, 29, 203-221. doi:10.1177/0265532211419826
- Hunt, K. (1965). *Grammatical Structures Written at Three Grade Levels*. Urbana, IL: National Council of Teachers of English.
- 稲木昭子・堀田知子・沖田知子. (2002). 『新えいご・エイゴ・英語学』. 東京：松柏社.
- 井之川睦美. (2010). 「教室内英作文と教室外英作文の比較—Error-free T-units にみられる正確さの分析を中心に」『外国語教育研究 外国語教育学会紀要』No.13 , 1-19.
- Intaraprawat, P. & Steffensen M. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language writing*, 4 , 51-69.
- 今井倫子・加藤直樹. (2014). 「メタ認知方略を使用した学習活動の効果（1）—高等学校英語学習のルーブリック開発」『岐阜大学教育学部研究報告 人文科学』第 63 巻第 1 号, 151-159.
- 石田知美. (2013). 「判別分析を使った学習者英語の言語的特徴」. 投野由紀夫・金子朝子・杉浦正利・和泉絵美(編). 『英語学習者コーパス活用ハンドブック』(pp. 100-107) . 東京：大修館書店.

- Ishikawa, S. (1995). Objective Measurement of low-proficiency EFL narrative writing. *Journal of Second Language Writing*, 4, 51-69.
- 石川慎一郎. (2005). 「大学生英語学習者の受容語彙と発表語彙力の関係ー語彙サイズテストおよびエッセイ・コーパス分析に基づくアプローチ」『中部地区英語教育学会紀要』第 34 号, 337-344.
- 石岡恒憲・橋本貴充・大津起夫. (2009). 「センター試験・英語と国語における素材文のリーダビリティと得点率についての統計的解析 (テキスト評価とリーダビリティ)」. 『電子情報通信学会技術報告. TL, 思考と言語』, 109(84), 1-6.
- Jacobs, G. (1986). Quickwriting: A technique for invention in writing. *ELT Journal*, 40, 282-290.
- Jacobs, H. L., Zinkgraf, S. A., Wormuch, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- Janoupoulis, M. (1992). University faculty tolerance of NS and NNS writing errors: A comparison. *Journal of Second Language Writing*, 1, 109-121.
- Johnson, K. & H. Johnson. eds. (1998). *Encyclopedia dictionary of applied linguistics*. Blackwell. (岡秀夫監訳. (1999). 『外国語教育学代辞典』. 東京: 大修館書店)
- Jones, S., & Tetroe, J. (1987). Composing in a second language. In A. Matsuhashi (Ed.), *Writing in real time: Modeling production processes* (pp. 34-57). Norwood, NJ: Ablex.
- Kaczmarek, Celeste, M. (1980). Scoring and rating essay talks. In *Research language testing*, John W. Oller, Jr. and Kyle Perkins (Eds.). New Burry House Publishers Inc. 151-159.
- 門田修平. (1990). 「作文評価における客観的指標の妥当性」『英語英文学新潮 1990 年版』ニューカレントインターナショナル. 339-350.
- 門田修平. (2002). 『英語の書きことばと話しことばはいかに関係しているか』東京: くろしお出版.
- 門田修平・池村大一郎編. (2006). 『英語語彙指導ハンドブック』. 東京: 大修館書店.
- 神尾昭雄・高見健一. (1998). 『談話と情報構造』. 東京: 研究社.
- Kanatani, et al. (1993). *The Role of Teacher Feedback in EFL Writing Instruction*. 平成 3 年ー平成 4 年度 文部省科学研究補助金研究.
- 加瀬正二郎. (1994) 「4 技能の実践的指導法」片山嘉雄. 遠藤栄一・佐々木昭・松村幹夫 (編) 『新・英語科教育の研究 (改訂版)』 (pp. 229-235). 東京: 大修館書店.
- Kochiyama, M., Aliponga, J., Johnston, C. C. (2013). Using rubrics in EFF classrooms: students perceptions and attitude. 『関西国際大学研究紀要』第 14 号, 41-47.
- 小池生夫編集主幹. (2003). 『応用言語学事典』. 東京: 研究社.

- 小池生夫・寺内 一・高田智子・松井順子・財団法人国際ビジネスコミュニケーション協会. (2010). 『企業が求める英語力 *English Skills: What Do Companies Really Need?*』. 東京：朝日出版社.
- 小島ますみ. (2011). 「英語学習者の産出語彙における語彙の豊かさ指標 S の提案と論証による S の妥当性」名古屋大学大学院国際開発研究科提出博士論文.
- 小島ますみ. (2013). 「語彙の豊かさと習熟度の関係」. 投野由紀夫・金子朝子・杉浦正利・和泉絵美 (編). 『英語学習者コーパス活用ハンドブック』 (pp.108-116). 東京：大修館書店.
- Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal* 25, 21-33.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19, 255-271.
- Laufer, B & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- Laufer, B & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- Laufer, B. & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: effects of language learning context. *Language Learning*, 48, 365-391.
- Lemmouh, Z. (2008). The relationship between grades and the lexical richness of students essays. *Nordic Journal of English Studies*, 7(3), 163-180.
- Linnarud, M. (1986). *Lexis in composition: a performance analysis of Swedish learners' written English*. Malmö : CWK GLEERUP.
- Malvern, D., & Richards, B. (1997). A new measure of lexical diversity. In A. Ryan, & A. Wray (Eds.), *Evolving models of language* (pp.58-71). Clevedon: Multilingual Matters.
- 松井恵美. (1979). 『英作文における日本人的誤り』. 東京：大修館書店.
- 松浪有・池上嘉彦・今井邦彦. (1983). 『大修館英語学事典』. 東京：大修館書店.
- McCarthy, P. M., & Jarvis, S. (2010). MTL, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods* 42 (2), 381-392.
- Meara, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16 (3), 5-19.
- 緑川日出子. (1994). 「ライティング」SLA 研究会 (編) 『第二言語習得に基づく最新の英語教育』. 東京：大修館書店. 287-299.

- 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的評価の関係」. 『統計数理研究所共同研究レポート 215 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』. 15 - 28.
- 水本篤. (2013). 「英文解析プログラムから得られる各種指標を使ったテキスト難易度の推定-教材作成への適用可能性-」. 『外国語教育メディア学会 (LET) 関西支部メソドロジー研究部会 2012 年度報告論集』, 142-150.
- 望月正道・相澤一美・投野由紀夫. (2003). 『英語語彙の指導マニュアル』東京:大修館書店.
- 望月昭彦. (2010). 「ライティング」望月昭彦・久保田章・磐崎弘貞・卯城祐司『新学習指導要領にもとづく英語科教育法』(pp. 149-160) 東京:大修館書店.
- 望月昭彦・深澤 真・印南 洋・小泉利恵 (編). (2015). 『英語 4 技能評価の理論と実践』東京:大修館書店.
- Mohan, B. A. & Lo, W. A. (1985). Academic writing and Chinese students: Transfer and developmental factors. *TESOL Quarterly*, 19, 515-534.
- Muncie, J. (2002). Process writing and vocabulary development: Comparing lexical profiles across drafts. *System*, 30, 225-235.
- Nadarajan, S. (2011). The challenges of getting L2 learners to use academic words in their writings. *Electronic Journal of Foreign Language Teaching*, 8(2), 184-200.
- Nakanishi, C. (2006). *A teaching approach to Japanese college students' EFL writing*. Keio University Press.
- 中島文雄・大塚高信. (1982). 『新英語学辞典』. 東京:研究社.
- Nation, I. S. P. (2013). *Learning vocabulary in another language (2nd ed.)*. Cambridge: Cambridge University Press.
- Negishi, M., Takada, T., & Tono, Y., (2012). A progress report on the development of the CEFR. *Studies in Language Testing*, 36, 137-157.
- Nishijima, H., Hayashi, K., Maseki, M., Kinshi, K., & Kuru, Y. (2007). Developing a writing rubric for classroom use in Japanese higher education. *Jacet Journal*, 45, 109-116.
- 縫部義憲. (1985). 「第3章7節 自由英作文」沖原勝昭(編). 『英語のライティング』(pp. 170-198) .東京:大修館書.
- 大井恭子. (2015). 「4 技能試験時代のライティングに必要な指導と評価とは」『英語教育』12月号. 大修館書店. 10-12.
- 沖原勝昭. (1994). 「英作文添削法と評価」『英語教育』12月号. 大修館書店. 29-31.
- 旺文社編. (2013). 『2014 年 受験用 全国大学入試問題正解 2 英語 (国公立大学編)』. 東京:旺文社.
- Perkins, K. (1983). On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *Tesol Quarterly*, 17,

- 651-671.
- Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, 47, 101-143.
- Polio, C. & Shea, M.C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10-27.
- Raimes, A. (1983). *Techniques in Teaching Writing*. Oxford University Press.
- Reid, J. (1990). Responding to different topic types: A qualitative analysis from a contrastive rhetoric perspective. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 191-210). Cambridge : Cambridge University Press.
- Richards, J. C., Platt, H., and Platt, J. (1992). *LONGMAN DICTIONARY of LANGUAGE TEACHING AND APPLIED LINGUISTICS*. Longman.
- 佐野正之. (1997). 「ライティング指導を見直すー和文英訳からパラグラフライティングまでー」『英語教育』12月号. 8-11.
- Sasaki, M. & Hirose, K. (1999). Development of an analytic rating scale for Japanese L1 writing. *Language Testing*, 16(4), 457-478.
- 染谷泰正. (2009). 「オンライン版「英語語彙難易度解析プログラム」(Word Level Checker) の概要とその応用可能性について」. 『青山学院大学文学部紀要』51, 97-120.
- Song, B. & Caruso, I. (1996). Do English and ESL faculty differ in evaluating the essays of native English-speaking and ESL students? *Journal of Second Language Writing*, 5, 163-182.
- Spaan, M. (1993). The effect of prompt in essay examinations. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing research*. Alexandria, VA:TESOL.
- 杉田由仁 (2015) 「ライティングの評価とテスト」『英語教育』1月号. 大修館書店. 18-19.
- 杉浦正利. (2008). 「英文ライティング能力の評価に寄与する言語的特徴について」. 成田真澄 (代表) 『学習者コーパスに基づく英語ライティング能力の評価法に関する研究』, 33-58. 平成 17 年度～平成 19 年度科学研究費補助金 (基礎研究 (C)) 研究成果報告書 (課題番号 17520394)
- 田畑光義・大井恭子. (2012). 「中学生へのパラグラフ・ライティング指導の効果」『関東甲信越英語教育学会 学会誌 KATE Journal』, 26 号, 79 - 91.
- 高梨庸雄・卯城祐司. (2000). 『英語リーディング事典』. 東京：研究社.
- 武田博. (1999). 「Fast writing と Comment feedback が日本人 EFL 中学生のエッセイに及ぼす影響について」『関東甲信越英語教育学会研究紀要』第 13 号. 1-14.
- 竹内理・水本篤. (2012). 『外国語教育研究ハンドブック 研究のより深い理解のために』東京：松伯社.

- 玉井健. (1993). 「談話的視点によるパラグラフィティングの効果的な指導」『中部地区英語教育学会紀要』第 22 号, 61-66.
- 田中敏・山際勇一郎. (1992). 『新訂ユーザのための教育・心理統計と実験計画法』東京：教育出版.
- Taylor, D. (2008). Readability. Info[Online] <http://www.readability.info/info.shtml>
- 寺澤芳雄主幹. (2002). 『英語学要語辞典』東京：研究社.
- 投野由紀夫. (2007). 『日本人中高生一万人の英語コーパス“JEFLL Corpus” 一中高生が書く英文の実態と分析一』東京：大修館書店.
- 投野由紀夫. (2013). 『CAN-DO リスト作成・活用 英語到達度指標 CEFR-J ガイドブック (CD-ROM 付)』東京：大修館書店.
- 植田一三・上田敏子・田岡千秋・小谷延良. (2015). 『CD BOOK TOEFL iBT TEST スピーキング+ライティング完全攻略』東京：明日香出版社.
- Van Deusen, B. (2014). An evaluation of a scoring rubric in the foreign language classroom. 『長崎国際大学論叢』第 14 卷, 69-76.
- Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17(1), 65-83.
- Vermeer, A. (2004). The relation between lexical richness and vocabulary size in Dutch L1 and L2 children. In P. Bogaards. & B. Laufer (Eds), *Vocabulary in a second language: Selection, acquisition and testing* (pp.173-189). Amsterdam and Philadelphia, PA: John Benjamins.
- West, M. (1953). *A general service list of English words*. London : Longmans, Green and Co.
- Wooddall, B.R. (2002). Language-switching: Using the first language while writing in a second language. *Journal of Second Language Writing*, 11. 7.-28.
- White, R. & Arndt, V. (1991). *Process Writing*. Longman.
- 山田 純. (1985) 「ライティングの能力・学力」沖原勝昭 (編) 『英語のライティング』 (pp. 44-81) . 東京：大修館書店.
- 安井稔・角谷裕子. (1998). 『英作文要覧』. 東京：開拓社.
- 米山朝二. (2002). 『英語教育実践から理論へ<改定増補版>』. 東京：松柏社.
- 米山朝二. (2003). 『英語教育指導法事典』. 東京：研究社.
- 吉島茂・大橋理枝 (他) 訳・編. (2004). 『外国語教育Ⅱ - 外国語の学習, 教授, 評価のためのヨーロッパ共通参照枠』. 東京：朝日出版社.
- 吉島茂. (2014). 「CEFR の日本の外国語教育への応用」『大学英語教育学会関東支部紀要』第 1 号, 4-19.
- Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly*, 16,

195-209.

Zamel, V. (1983). The Composing Process of Advanced ESL Students: Six Case Studies.

TESOL Quarterly, 17, 165-187.