

多様な英語能力の測定

Measuring Different Aspects of Language Proficiency

原田 康也¹, 鈴木 正紀², 森下 美和³
Yasunari HARADA, Masanori SUZUKI, Miwa MORISHITA

¹早稲田大学, ² Pearson, ³神戸学院大学
Waseda University, Pearson, Kobe Gakuin University
harada@waseda.jp, masanori.suzuki@pearson.com, miwa@gc.kobegakuin.ac.jp

Abstract

Japan is faced with an imminent challenge of cultivating ‘global human resources’, as the whole society develops into a global information society. The course of study defined by the Ministry of Education, Culture, and Sports and Technology (MEXT) has emphasized ‘communication skills’ as a focus of the foreign language subjects (e.g., English) since the 1990’s, with additional emphasis on cultivating “attitudes toward active engagement in communication” across the board in all subjects since around the turn of the century and the millennium. English listening component was incorporated into the University Entrance Examination Center Tests in 2006 but how its backwash affected listening proficiency of Japanese college students in general remains questionable. During the academic year of 2014, the Educational Reform Working Group within the Liberal Democratic Party suggested use of TOEFL iBT as part of university entrance procedures and in 2015, MEXT advised universities to consider employment of 4-skills English language tests in admission processes. However, how students would perform in those tests is not known as speaking or writing tests have not been administered in large scales in Japan.

In the school years of 2015-2016 and 2016-2017, approximately 60 and 90 first-year students, respectively, studying at a private university in Tokyo took Versant English Test, an automated test of spoken English, and Versant Writing Test, an automated test of reading and writing English, as well as Oxford Quick Placement Test, which measures vocabulary, collocation, and grammar through multiple choice reading tasks. The present study reports results and analyses of the test scores and estimated CEFR levels, and then discusses challenges that Japanese learners and teachers of English are trying to deal with.

Keywords — Language Proficiency, Automated Spoken Test, Automated Writing Test, CEFR

1. 大学入学者の英語力の多様性

国際情報社会の本格的な進展にともない、『グローバル人材開発』が緊急の課題として認識されるようになった。文部省・文科省の定める学習指導要領を見ると1990年代以降は外国語科目（英語）の授業において「コミュニケーション」が重視されるようになり、2006年度から大学入試センター試験に英語リスニングが導入され、2011年度より小学校5-6年生で英語が必修となり、2012年度より高校での英語の授業を原則として英語で実施するように指導要領に明記されるようになった。このように、日

本人英語学習者の英語口頭運用能力の向上を目指した一連の改革が進められ、一定の変化・改善が見られてもおかしくない状況にあるが、中学・高校における英語学習・英語指導の中核が大学入試に向けての訳読式学習にあり、口頭英語の運用能力の向上は見られないという感想も根強い。このため、2014年度に自民党の教育再生実行本部が大学入試に TOEFL 導入を提言し、2015年度には文科省が「高大接続」と「大学入試改革」に先行して4技能英語外部試験を活用することを大学に求めた。大学入試センター試験には毎年50万人規模の受験者があり、現在提唱されている「大学入学共通テスト（仮称）」は少なくとも同程度の受験者数となることが見込まれるため、記述式問題に対する大量の回答を短期間に処理する採点システムの開発が期待されるが、具体的な報告はない。

大学入学者の英語力は多様である。対象を主に日本国内の、あるいは日本の制度に基づく高等学校を卒業し、一般的な学力試験に基づく入学試験を経て入学してきた学生に限定しても、中学・高校と英語が比較的得意で熱心に勉強してきた学生もいれば、英語が苦手な最低限の勉強しかしてこなかった学生もいる。多くの私立大学の入学試験において、受験する科目が3教科3科目以下である中でも英語が入試科目に入ることが多いかもしれないが、複数科目の合計点が合否の基礎となるため、英語を得点源として合格を果たす受験生がいる一方で、英語の得点が振るわない中で他教科・他科目の得点が良いために合格する受験生も多く、また入学後の学業成績や就職状況に対してどちらがより強く相関するかも一概には結論付けることが難しい状況となっている。

一般的な入学試験のほかに、付属高校などからの推薦入学者・自己推薦や学校推薦などの推薦制度による入学者、いわゆるAO方式による入学者など、英語の学力を直接的に問われることなく大学に入学する学生も増えており、この中には中学・高校時代に学ぶべき英語の語彙・文法・4技能の運用能力を十分に身に着けないまま入学してくる学生もいる。他方、いわゆる帰国生向けの入学試験を通じて入学する学生に限らず、英語圏または英語を媒介言語とする学校での学習を複数年または長期にわたって経験してきた学生も増えている。

早稲田大学法学部の新入生を例にとると、同一のクラスの中に、TOEIC スコアでいうと 300 点台あるいはそれ以下の学習到達度の学生もいる一方で、900 点を超える学生もいるというような、学習到達度に大きな幅が見られる。短期または複数年にわたる海外・英語圏での学修経験はもはや例外とは言えず、いわゆる帰国生向けの入学試験による入学者の数は限られているが、一般の入試や推薦を通じて入学する学生も含め、クラスによっては 1 割から 3 割ぐらいがこうした学生となっている。

中学・高校を日本の制度の中で過ごしてきた学生の英語学習経験も、今日では極めて複雑な状況となっている。小学校での英語学習経験が個別的・多層的になってきている一方で、学校外の英語塾・英会話学校などでの学修経験も、保護者の意向や地域的な特性も含め、学生の英語運用能力に大きな影響を与えている。中学・高校における教科としての英語学習経験も、コミュニケーション活動・リスニング・スピーキング・ライティングを重視する学校もあれば、旧来の入学試験対策を重視する学校も多い。そうした中で、同じく TOEIC が 400 点・500 点・600 点の学生であっても、その英語 4 技能を見てみると、学生ごとに大きなばらつきがある。

これとは別に、日本の大学で英語を学ぶ学生の英語力が多くの場合に初歩というには学習段階が進んでおり、一方で一部の例外を除くと上級に達していないという意味において、ざっくりと中級段階にあることにも着目する必要がある。初級の学習者は、英語に関して習得した語彙・文法などの知識も限定的で、運用についても押しなべて限られたスキルとなっている。また、スピーキングやライティングなどの産出スキルにおいても、語彙が限られるためにおおむね短い文・短い発話に限られ、したがってその内容も概して単純となる。一方、上級者は豊富な語彙・文法を習得済みで、どのような話題についてもある程度の対応ができ、スピーキングにおいてもライティングにおいても、安定した能力を示す。これに比べると、「中級」というのは、一様な「中級さ」を示すものではなく、スピーキングについてはある程度まで熟達しているのにライティングは経験不足、あるいはその逆というような状態であり、語彙・文法の知識はそれなりに獲得しているのに、運用面でこれを活用する経験と訓練が不足している状態を指す。スピーキング・ライティングなどの産出においても・リーディング・リスニングなどの受容においても、ある種の話題については個人的な興味関心から豊かな内容を扱うことができるのに対して、興味の薄い話題に関しては全く初級レベルの反応しか示すことができないなど、話題・分野によるムラが大きく、また、書く文・話す発話も、短かったり長かった

り、複雑な内容を含んでいたり、きわめて単純であったりと揺れ幅が大きいということが中級の特徴である。

このような点を考慮すると、日本人大学生の英語力を語彙・文法などの言語知識と 4 技能を含めて総合的に考える上で、一つの全体的なスコアで学力を見ようとすることに無理があることが納得できるであろう。本稿で詳述するように、筆者が授業を担当する早稲田大学法学部の学生を見た場合、第一に目に付くのが、語彙・文法・連語などの英語に関する知識とリスニング・スピーキング・ライティングなどのリアルタイム性を要求されるタスクや産出を中心とした課題における運用能力に大きな乖離がある点である。また、同じく産出といっても、スピーキングとライティングを取ると、そこには一定の差が見られる。こうした点は、授業を担当する教員の実感としてはこれまでも繰り返し報告されてきたことではあるが、比較可能な試験のスコアとして明確に示すような報告はこれまで必ずしも多くなかった。

このような、知識・技能の乖離とスキルごとの習熟度の差を前提とすると、現在多くの大学で試みられているスキル別クラス編成と事前テストによるクラス分けも全体としての整合性に疑問が生じる。例えば、ある大学のある学部でスピーキングあるいはライティングを重視したクラスを編成し、さらにこれを習熟度別にクラス分けしようとした場合、理想的にはスピーキングあるいはライティングなどそれぞれの授業で中心となるスキルの習熟度に応じたクラス分けを行うべきであると考えるのが自然であるが、実際には時間的な制約などから TOEFL IP / TOEIC LR ITP / TOEIC Bridge / CASEC など、4 技能ではなくリーディングとリスニングを中心とした課題の中で多肢選択によって文法や語彙などの知識を測る試験を用いることになりがちである。こうした試験で同様のスコアを得た学生のスピーキングやライティングの熟達度が同様であるという根拠は多くの場合に存在しない。あるいは、大学入学者のスピーキングやライティングの熟達度が低いことから、4 技能試験を入学者選抜の過程に組み入れたとしても、総合点や平均点を使う限り、受験者としては比較的自信のあるリーディングやリスニングで総合的な得点を向上させようとする可能性も高く、本来の狙いであるスピーキングやライティング能力の向上につながらない可能性も懸念される。

2. Versant Tests と OQPT

Versant English Test (VET) は外国語・第二言語としての英語の口頭運用能力を判定する全自動試験で、試験の開発設計は人間の言語処理の現実を踏まえており、極

めて高い信頼性・妥当性・運用性を示す。Versant Writing Test (VWT) は書記言語での英語運用能力を測定する全自動試験である。

VET は表 1. に示すように全 6 設問で構成される。指示文と問題は全て電話もしくはコンピュータ上で配信され、回答は音声処理システムを用いて自動採点され、面接官や評定者は必要としない。VET ではリアルタイムで行われる自然な会話場面において中核になるとされる心理言語学的な言語処理・産出能力を測定している。英語の話し言葉コーパス Switchboard Corpus の上位 8000 語に準拠してテスト項目を作成し、英語母語話者により自然な会話の速度で録音されているため、大多数の日本人大学生にはそれほど簡単ではないことがこれまでの試行から明らかとなっている。

表 1. Versant English Test の構成

設問	タスク
Part A	Reading (音読)
Part B	Sentence Repeat (復唱)
Part C	Short Answer Questions (語句回答)
Part D	Sentence Builds (文の構築)
Part E	Story Retelling (話の再現)
Part F	Open Questions (自由回答)

Part A: Reading では試験用紙に印刷されている文の音読を求められる。音声により指示された番号の文を読み上げる。Part B から Part F の問題は全て音声のみで出題され、試験用紙には各設問の指示文と例題だけが印刷されている。なお、Part A から Part E までの設問が自動採点の対象となり、総合点と構文・語彙・流暢さ・発音の正確さの 4 つのサブスキルが 20-80 点で報告される。

Versant English Test の妥当性検証の研究結果によると、総合点におけるテストの信頼性は split-half 法で 0.97 が、test-retest 法でも 0.97 が報告されており、高い信頼性が実証されている。併存的妥当性の検証結果としては、TOEFL-iBT のスピーキングセクションの点数と 0.75 (n=130)、IELTS のスピーキング試験と 0.76 (n=130)、TOEIC リスニングセクションと 0.71 (n=171) など、著名な他の英語試験とも比較的強い相関がある。詳細は [9]・[10] を参照されたい。

VWT は表 2. に示すように全 5 設問で構成される。指示文と問題は音声・文字でインターネットを通じてコンピュータのブラウザ上に配信され、キーボードからタイプする回答は自動採点されるため、面接官や評定者を必要としない。VWT の設問には一部ビジネス場면을前提とした email の作成課題があるため、日本人大学生とつ

て内容的な困難がある可能性が懸念されたが、これまでの試行ではこの点は大きな障害となっていない。むしろ、英文で示される設問の指示を十分に理解しないまま回答する受験者がたまにいることの影響が大きいかもしれない。

表 2. Versant Writing Test の構成

設問	タスク
Part A	Typing (タイピング)
Part B	Sentence Completion (空所補助)
Part C	Dictation (ディクテーション)
Part D	Passage Reconstruction (文章の再構築)
Part E	E-Mail Writing (Eメール作成)

総合点と文法・語彙・構成・文体・読解の 5 つのサブスキルが 20-80 点で算出・報告される。Part A の Typing からはタイプ速度とタイピングの正確さが算出されるが、これらはテストの総合点には含まれず、受験者に関する追加情報として提示される。VWT はコンピュータ受験でありキーボードを使用して回答入力を行うため、受験者がタイピングに不慣れでタイプ速度が極端に遅い、あるいは正確さが低い場合には、テスト結果がライティング力よりもタイピング力の欠如を反映している可能性があるため、タイプ速度とタイピングの正確さが追加情報としてスコアレポートに含まれている。

VWT の妥当性検証の研究結果によると、総合点におけるテストの信頼性は split-half 法で 0.98 が、また人為採点によるスコアと自動採点によるスコアとの間には 0.98 の相関が報告されており、VET と同様に高い信頼性が実証されている。詳細は [11] を参照されたい。

Oxford Quick Placement Test (OQPT) には paper and pencil version と CD-ROM version と online version があるが、本稿で報告するデータは実施が簡便な paper and pencil version による。英語の読解力および語彙・文法の知識を問う問題が全部で 60 問あり、所要時間は 30 分である。Part 1 の 40 問はすべての学生が対象(中級以下のレベル)で、Part 1 のみを使用する場合の所要時間は 20 分である。Part 2 の 20 問は Part 1 の 40 問中 36 問以上に正答した上級レベルの学生に使用することが推奨されている。詳細は [12] を参照されたい。今回はさまざまな習熟度の学生を比較するため、上記の詳細に関わらず、Part 1 と Part 2 全 60 問の合計得点を使用した。

VET・VWT・OQPT の各テストでは、テストの点数とそこから推定される受験者のヨーロッパ言語共通参照枠組(Common European Framework of Reference,

CEFR) の各レベルとへの対応付けが公開されている。

CEFR では実践的な言語使用場面を想定し、各レベルの記述に機能的 (functional) な要素と場面的 (situational) な要素を含む Can-Do Statements を使っている。具体的には表3にある Can-Do Statements にて各レベルが定義されている。詳細は [13] を参照されたい。

表3. CEFR の A1 から B2 レベルの能力記述抜粋

A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him / herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.
A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.

後に示すように、VET の結果によると受験した学生の大半は A レベルとなるが、OQPT の結果からは大部分の学生のレベルが CEFR の B レベルに相応する。A レベルは「Basic User」で B レベルは「Independent User」と

広く定義されている。表3 で目につく表現としては familiar / immediate relevance / need / simple / routine などがある。A レベルの言語使用者は、個人的な身近な話題について頻繁に使用される基礎的な表現を理解・使用することができるが、複雑で抽象的な表現を理解したり発話したりすることはできない。VET のスコアから分かることは、大半の学生が A レベルであり、A1 レベルでは、一般的な会話速度では言語処理できず、話し相手にゆっくりと明瞭に話してもらうなどの対応が必要となる。

表4. VET・VWT・OQPT のスコアと CEFR レベルの対応表

CEFR	VET	VWT	OQPT
>A1	20-25		0-9
A1	26-35	20-29	10-17
A2	36-46	30-43	18-29
B1	47-57	44-53	30-39
B2	58-68	54-66	40-47
C1	69-78	67-76	48-54
C2	79-80	77-80	55-60

なお、VWT は設問の詳細指示は英語で示され、またキーボードからの回答が前提となるため、中学・高校でキーボードを使って英作文を行う経験のほとんどない日本人英語学習者にとっては、大学生といえどもコンピュータ教室などでの英語学習経験をある程度積んだ段階でないと、意味のあるスコアが得られない。

3. 被験者と受験時期

2004 年度以降に早稲田大学法学部に入学する新入生は 800 名前後であるが、データ収集の対象となる新入生はこのうち 3 クラス 60 名弱から 90 名強であり、人数として 1 割前後、1 年生全体から無作為抽出などの統計的分析にふさわしい方法で選別されたわけではなく、たまたま特定のクラスに割り振られたということであるため、その代表性には若干の疑念が残る。しかし、より代表性の高いこれを代替するデータは存在しない。また、少なくとも VET については 2000 年度から 2017 年度まで、分散以外はほぼ同様な数字を示しているため、法学部の新入生全般の英語運用能力について知る一つの指標として十分な意味があると考えている。(分散については、特定少数の学生が極めて高いスコアを得たときに大きな値となるため、年度によって変動が見られる)

2015 年度には早稲田大学法学部 1 年生に在籍する 62 名が参加した。実施期間は 2015 年度の年間を通して行い、VET と VWT は各 4 回、OQPT は 3 回実施した。VET

については 2015 年度春秋学期開始・終了に合わせて 4 月初旬・7 月上旬・9 月下旬・1 月上旬の授業時に試験用紙を配布し、できるだけ速やかに受験することという指示を出し、ほぼ全員が2週間以内に受験を完了した。VWT についてはほぼ同時期に（春学期初めは学生のコンピュータへの習熟を考慮し少し遅れて）授業時に実施した。OQPT については、連休前の 4 月最後の授業で授業時間内に 30 分を使って実施し、7 月、12 月にも同様に実施した。OQPT には同レベルの Version 1 と Version 2 があり、1 回目は前者を、2 回目と 3 回目は後者を使用した。全 62 名の被験者中、VET と VWT を全て予定通りに受験し、スコアが算出された学生は 53 名であった。この 53 人は OQPT もほぼ全員が全 3 回を受験したため（4 月 52 名、7 月 53 名、12 月 50 名）、以下の分析ではこの 53 人を対象としている。

同様に 2016 年度には 91 名の 1 年生のうち、74 名についてスコアが整っており、OQPT も全員が全 2 回を受験したため、以下の分析ではこの 74 名を 2016 年度のデータ対象とする。（2013 年度から 2016 年度までのデータにより、OQPT についても大体の傾向は明らかになっている。また、VWT は 2015 年度に初めて実施したため、教室での実施に際して授業担当者が若干不慣れなところもあり、実施時間が 60 分近くと一回の授業時間の大部分を使ってしまったため OQPT の実施回数を減らして対応した）

4. データ分析

4.1. 記述統計量

VET、VWT、OQPT の総合点の記述統計量を受験回毎に表 5 から表 10 にまとめる。

表 5. 2015 年度 VET 受験回別記述統計量(n=53)

	4 月	7 月	9 月	12 月
平均	38.89	41.47	40.57	42.42
標準分散	6.77	6.45	6.18	5.54
最低	27	29	25	31
最高	55	62	56	60

表 6. 2016 年度 VET 受験回別記述統計量(n=74)

	4 月	7 月	9 月	12 月
平均	38.16	40.49	39.42	41.11
標準分散	5.39	5.66	6.28	5.62
最低	27	29	27	32
最高	62	59	63	62

表 7. 2015 年度 VWT 受験回別記述統計量(n=53)

	4 月	7 月	9 月	12 月
平均	41.45	44.08	44.51	47.75
標準分散	7.32	8.17	8.00	7.94
最低	20	20	22	22
最高	62	66	64	67

表 8. 2016 年度 VWT 受験回別記述統計量(n=74)

	4 月	7 月	9 月	12 月
平均	41.24	43.55	44.04	45.70
標準分散	6.79	6.78	7.12	6.78
最低	26	24	25	28
最高	58	62	58	60

表 9. 2015 年度 OQPT 受験回別記述統計量

	4 月 (n=52)	7 月 (n=53)	-	12 月 (n=50)
平均	37.08	38.06	-	38.24
標準分散	6.57	5.98	-	6.44
最低	20	25	-	21
最高	50	48	-	50

表 10. 2016 年度 OQPT 受験回別記述統計量

	4 月	7 月	-	-
平均	36.28	37.70	-	-
標準分散	6.07	6.36	-	-
最低	20	21	-	-
最高	50	50	-	-

VET・VWT の平均点を年度初めの 4 月と年度末の 12 月で比較すると、向上していることが明瞭である。対応ありの t 検定をそれぞれ行った結果、統計的な有意差 ($p < 0.05$) が見られた。一方 OQPT の平均を 4 月と 12 月で比較すると、若干向上しているものの、それほど大きな差はなく、対応ありの t 検定を行った結果、統計的な有意差は見られなかった ($p > 0.05$)。

上記の記述統計量のまとめからすると、比較的少数の学生のスコアから一般的な傾向を述べようとしているように見えるかもしれないが、筆者たちは 2000 年度から毎年法学部新入生のうち 50 人から 100 人弱程度のデータを取得しており、少なくとも Versant English Score のスコアについて、その傾向に大きな変動がないことを見てきている。2006 年度から 2011 年度までの Versant English Test の受験者数と総合点の記述統計量の一部を年度別にまとめたのが次の表 11 である。（鈴木・原田による一連の報告 [1]・[2]・[4] を参照されたい。）

表 11. 年度別記述統計量

年度	N	平均	標準 偏差	分散	最 高	最 低
2006	90	37.27	6.95	48.31	62	21
2007	74	37.26	8.27	68.39	66	23
2008	74	36.01	6.43	41.41	51	23
2009	71	37.11	5.90	34.82	49	25
2010	86	37.30	7.47	55.86	80	23
2011	97	37.43	6.00	36.02	53	20

平均点は各年度 36 点・37 点前後であり、大きな変動がないことが見て取れる。一元配置分散分析による統計分析でも、平均点の差には統計的な有意差は見られなかった。[F(5,486) = 0.45, n.s.]

4.2. CEFR からの考察

各テストのスコアから CEFR レベルに換算し、受験月別にどのような分布になるか、表 12 から表 17 にまとめた。

表 12. 2015 年度 VET から CEFR レベルへ換算 (n=53)

CEFR	4 月	7 月	9 月	12 月
A1	19 (35.8%)	8 (17.0%)	10 (18.9%)	6 (11.3%)
A2	27 (50.9%)	36 (67.9%)	36 (67.9%)	35 (66.0%)
B1	7 (13.2%)	6 (11.3%)	7 (13.2%)	11 (20.8%)
B2	0 (0%)	2 (3.8%)	0 (0%)	0 (0%)
C1	0 (0%)	0 (0%)	0 (0%)	0 (0%)
C2	0 (0%)	0 (0%)	0 (0%)	0 (0%)

表 13. 2016 年度 VET から CEFR レベルへ換算 (n=74)

CEFR	4 月	7 月	9 月	12 月
A1	25 (33.8%)	19 (13.5%)	18 (24.3%)	7 (9.5%)
A2	44 (59.5%)	56 (75.7%)	48 (64.9%)	57 (77.0%)
B1	4 (5.4%)	6 (8.1%)	7 (9.5%)	9 (12.2%)
B2	1 (1.4%)	2 (2.7%)	1 (1.4%)	1 (1.4%)
C1	0 (0%)	0 (0%)	0 (0%)	0 (0%)
C2	0 (0%)	0 (0%)	0 (0%)	0 (0%)

表 14. 2015 年度 VWT から CEFR レベルへ換算 (n=53)

CEFR	4 月	7 月	9 月	12 月
A1	3 (5.7%)	1 (1.9%)	2 (3.8%)	2 (3.8%)
A2	32 (60.4%)	22 (41.5%)	21 (39.6%)	10 (18.9%)
B1	16 (30.2%)	25 (47.2%)	23 (43.4%)	30 (56.6%)
B2	2 (3.8%)	5 (9.4%)	7 (13.2%)	10 (18.9%)
C1	0 (0%)	0 (0%)	0 (0%)	1 (1.9%)
C2	0 (0%)	0 (0%)	0 (0%)	0 (0%)

表 15. 2016 年度 VWT から CEFR レベルへ換算 (n=74)

CEFR	4 月	7 月	9 月	12 月
A1	3 (4.1%)	1 (1.4%)	4 (5.4%)	1 (1.4%)
A2	41 (55.4%)	38 (51.4%)	27 (36.5%)	26 (35.1%)
B1	28 (37.8%)	28 (37.8%)	35 (47.3%)	39 (52.7%)
B2	2 (2.7%)	7 (9.5%)	8 (10.8%)	8 (10.8%)
C1	0 (0%)	0 (0%)	0 (0%)	0 (0%)
C2	0 (0%)	0 (0%)	0 (0%)	0 (0%)

表 16. 2015 年度 OQPT から CEFR レベルへ換算

CEFR	4 月 (n=52)	7 月 (n=53)	12 月 (n=50)
A1	0 (0%)	0 (0%)	0 (0%)
A2	7 (13.5%)	4 (7.5%)	3 (6.0%)
B1	25 (48.1%)	25 (47.2%)	27 (54.0%)
B2	19 (36.5%)	21 (39.6%)	16 (32.0%)
C1	1 (1.9%)	3 (5.7%)	4 (8.0%)
C2	0 (0%)	0 (0%)	0 (0%)

表 16. 2016 年度 OQPT から CEFR レベルへ換算

CEFR	4 月	7 月
A1	0 (0%)	0 (0%)
A2	14 (18.9%)	9 (12.2%)
B1	40 (54.1%)	32 (43.2%)
B2	18 (24.3%)	31 (41.9%)
C1	2 (2.7%)	2 (2.7%)
C2	0 (0%)	0 (0%)

2015年度のスコアを例にとると、リスニングとスピーキング力を測定する VET では、大多数 (85%程度) の学生が A1 と A2 レベルにあるが、リーディングとライティング力を測定する VWT では大多数が A2 と B1 レベルになる。回答が選択式であるため、言語産出が必要とされない OQPT では、学生は B1 と B2 レベルにあると判定される。さらに VET での 4 月と 12 月の CEFR レベルの分布を比較すると、4 月では A1 と A2 で全体の 86.7% を占めていたが、12 月では A2 と B1 が 86.8% を占めるようになり、分布がシフト (点数が向上) したことが観察される。詳細な数字はさておき、全体の傾向として同様の年度内の変化が毎年見られている。

VWT を見てみると、2015 年 4 月時点での分布は A2 と B1 が 90.6% を占めていたが、そのうちの 60% 以上が A2 だった。しかし、12 月になると B1 レベルの学生数が 56.6% まで増加した。こちらも VET と同様に分布がより上位のレベルへとシフトしたことが明確である。

これに比して、OQPT の結果からはこのような年度内の分布の変化は見られず、データ分布は比較的一貫していた。OQPT では 4 月から C1 レベルの学生が 1 名であるが存在し、C1 の学生数が若干の増加傾向にあった。

5. 考察

VET・VWT・OQPT では各テストが英語能力試験ではあるものの、テストの構成概念や評価対象のスキルなどが異なる。本研究がこのように 3 つの異なるテストを用いることで示唆されるところはいくつかある。

「日本人英語学習者はリスニングやスピーキングが弱く、リーディングやライティングは比較的良い」と一般的に言われるが、本研究のデータは、相対的なレベルについてある程度これを反映している。「知識型のマークテストでは得点が取れるが、実際に英語を使用することはうまくできない」という相対的な比較も反映している。

鳥飼は [2] で TOEIC と TOEFL のスコアを中心に、時系列的・国際的な比較を試みている。(受験者が英語学習者のどの層に対応するか配慮してスコアを解釈する必要があることを忘れてはならないが、) TOEFL のスコアに関して、鳥飼は「日本はリスニングもよくはないが、文法もダメ、読むのも弱い。」(p.99) と「読み書きはできるが会話はだめ」というありがちな誤解を正した上で、『実践的コミュニケーション能力』育成に大きく指導方針が転換し、限られた授業時間内に聞くこと、話すことの指導に重点が置かれたのであるから、結果として「読む力」が相対的に弱くなったのは理解できるとしても、それを補うだけの『リスニング力』が高まっているわけ

ではないのはなぜなのか。」(p.107) と 2000 年前後の中学・高校の英語教育の課題を指摘している。

大学で英語を担当する教員の実感としては、2017 年においても、中学・高校の英語学習段階で口頭表現やコミュニケーションの訓練どころか、教科書の内容に即した聞き取り練習やテキストの音読など、音声面の基礎練習すら十分に実施されていないのではないかという印象が払拭できず、新入生の (語彙・文法・読解などの) 基礎学力は (他の教科と同様に) 低下しているという印象が強い。残念ながら、日本における学力テストの多くは、偏差値に代表されるように母集団内の相対的評価を目的とするので、この点を裏付ける、あるいは否定するような客観的で広範囲にわたる時系列的なデータは、広く公開される形で国内に存在しない。

知識型の OQPT で学生が B1 と B2 レベルであることは、対象となる学生がいわゆる「偏差値」が比較的高い大学・学部に所属していることと、入学試験において英語が一定の割合を占めている点から、当然予想される結果ではあるが、リスニング・スピーキングやリーディング・ライティングの運用能力に比較して知識のレベルが高いといっても決して C1 ないし C2 レベルの学生が大多数というわけではなく、英語知識の学習到達度としても「仕事で英語が使える」まで高いレベルではない。英語運用型試験となる VET や VWT では、その知識すら有効に利用できていないことがわかる。リアルタイムでの処理が必要とされる VET では大多数の学生が CEFR の定義によると A1 と A2 レベルであり、CEFR による「Basic User」となり、長年の英語教育の課題の一つである「使える英語」の達成への道のりは遠い。(知識と運用の乖離については [5] で報告し、[6] で対応策について検討したうえで [7] で日本人英語学習者の wh-疑問文に対する知識と習熟度が低い点を指摘している。)

VET に比べて VWT のスコアが高く、CEFR で 1 レベルほど上と判定される理由はまだ明らかになっていないが、1990 年代から東大二次試験を嚆矢として国立大学に広まった英語「自由英作文」の washback として捉えることができるかもしれない。国立大学受験を意識する高校生は、対応策の徹底度は異なるとはいえ、入試当日の限られた時間の中で 1 パラグラフから数パラグラフの回答を英文でまとめることを求められるようになったため、1980 年代までに比べると 2010 年代になってライティングについてある程度の向上が見られたとしても不思議ではない。

年間でのスコア向上も VET については標準測定誤差をかるうじて超える程度であるのに対して、VWT についてはある程度明確な上昇が見られる。対象となった学生

は授業の中で年間を通じて比較的集中的な口頭での発話（質問に対する応答）と作文の訓練を授業の中で繰り返している。90分の授業時間の中で30分弱の応答練習（3人ずつのグループで質問に対して回答する練習）と20分弱の小グループでの発表と質疑応答の練習を毎週繰り返すとはいえ、冷静に一人が英語の産出に用いる授業内の時間を計算すると、応答練習で1回の発話時間は45秒と想定しており、30分の応答練習で一人が発話する時間はせいぜい3分程度である。少人数でプレゼンテーションと質疑応答を行う練習でも、一人当たりの発話時間はどのように多めに見積もっても3分に満たない。一方、作文に費やす時間は、当初は90分の授業の中で40分程度、その後だんだんと切り詰めていくが、隔週で授業内に20分から30分近い時間を費やし、宿題としては1時間から6時間ぐらいを費やす学生がいるため、授業外に英語で発話する時間がほとんどないと見込まれる点と比較すると、年間の産出時間においてスピーキングとライティングでは10倍から100倍の違いがあることになる。

話し言葉であれ、書き言葉であれ、英語をどれだけ容易にリアルタイムで処理・産出・運用していけるようになるかが「英語教育」と「グローバル人材育成」の課題であると捉え、どのように英語運用面において次のレベル（Bレベル: Independent User）に導いていけるかが大学英語教育の急務である。

6. 謝辞

本稿で紹介する早稲田大学法学部1年生からのデータの収集並びに分析にあたっては、早稲田大学特定課題研究助成費（特定課題B）課題番号2016B-022『日本人英語学習者の疑問文産出能力：現状調査と形態統語処理自動化による運用能力向上』（研究代表者：原田康也）・早稲田大学特定課題研究助成費（基礎助成）：課題番号2016K-028『英語疑問文構築のための形態統語処理能力の現状調査と向上訓練』（研究代表者：原田康也）・科研費基盤研究(C)：課題番号16K02946『英語コミュニケーションにおける統語的プライミングを利用した統語処理の自動化促進』（研究代表者：森下美和）・科研費基盤研究(B)：課題番号15H03226『日本人英語学習者のインタラクション（相互行為）を通じた自律的相互学習プロセス解明』（研究代表者：原田康也）・科研費基盤研究(A)：課題番号26244031『学習による気づき・注意機能および相互的同調機能と第二言語情報処理の自動化プロセス』（研究代表者：横川博一）・科研費基盤研究(B)：課題番号26285163『外国語力と外国語副作用の関係：外

国語力が等しければ思考力の低下量も等しいか？』（研究代表者：高野陽太郎）・早稲田大学特定課題研究助成費（基礎助成）：課題番号2015K-025『言語処理短期記憶の強化による英語処理の自動化と運用能力向上の実証』（研究代表者：原田康也）の助成を得ている。

参考文献

- [1] 鈴木正紀・原田康也, 「大学新入生の英語リスニング・スピーキング熟達度の定点観測」, 電子情報通信学会技術報告, vol. 111, No. 320, pp. 37-42, 2011年11月19日.
- [2] 鈴木正紀・森下美和・原田康也, 「言語技術の言語評価への応用: 多様な英語能力の測定」, 電子情報通信学会技術報告, vol. 116, No. 77, pp. 41-46, 2016年5月28日.
- [3] 鳥飼久美子, 「TOEFL・TOEICと日本人の英語力」, 講談社現代新書 1605, 講談社, 2002年4月20日.
- [4] 原田康也・鈴木正紀, 「大学新入生の英語リスニング・スピーキング熟達度の定点観測に向けて」, 電子情報通信学会技術報告, vol. 110, No. 63, pp. 21-25, 2010年5月21日.
- [5] 原田康也・森下美和, 「日本人英語学習者の言語処理と言語運用能力: Versant English Test のスコアを中心に」, 信学技報, vol. 113, No. 174, pp. 1-6, 社団法人電子情報通信学会, 2013年7月27日.
- [6] 原田康也・森下美和, 「大学英語教育における知識と運用の統合: 文法知識の運用課題と実体的コミュニケーションの場の提供」, 電子情報通信学会技術報告, vol. 114, No. 385, pp. 19-24, 2014年12月13日.
- [7] 森下美和・原田康也, 「日本人英語学習者のwh疑問文の知識と運用に関する調査: 習熟度の観点から」, 日本認知科学会第32回大会発表論文集, pp. 947-951, 日本認知科学会, 2015年9月18日.
- [8] 文部科学省, 高大接続システム改革会議「最終報告」, 2016年3月31日.
http://www.mext.go.jp/b_menu/shingi/chousa/shougai/03/3/toushin/_icsFiles/afieldfile/2016/03/31/1369233_01_2_1.pdf [最終アクセス 2016年5月12日]
- [9] Pearson, Versant English Test – Test Description and Validation Summary, Menlo Park, CA, 2013.
- [10] Pearson, Versant English Test Can-Do Guide, Palo Alto, CA, 2003.
- [11] Pearson, Versant Writing Test – Test Description and Validation Summary, Menlo Park, CA, 2013.
- [12] Oxford University Press. (2004). Quick placement test. Oxford: Oxford University Press.
- [13] Council of Europe, Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), 2001.
http://www.coe.int/t/dg4/linguistic/Source/Framework_E_N.pdf [last accessed 2014/06/03]