聖徳大学研究紀要　聖徳大学　第30号　聖徳大学短期大学部　第52号　103-110（2019）

TOEIC と語彙学習教材との関係性におけるコーパス分析
－語彙の多様性と複雑さの比較に焦点を当てて－

長橋 雅俊

Corpus Analysis of the Relation between TOEIC and Wordbooks
－ Focusing on Comparison of Lexical Varieties and Complexities －

NAGAHASHI, Masatoshi

要旨

日本の大学では、今や学生に向けた国際コミュニケーション英語能力テスト（Test of English for International Communication, 以下 TOEIC）の受験が推奨され、スコアに応じて単位認定などの優遇措置を与える学校が増えつつある。また、就職活動における応募要件の一部に TOEIC スコアを求めたり、社員の昇進条件として課す企業も少なくない。しかしながら、IIBC（2019）の報告によれば、2018 年の日本人受験者の平均スコアは 520 点であり、他のアジア諸国より低い状況が続いている。本研究では、延べ 12 セットの TOEIC 公式サンプル・テストと計 5 冊の語彙学習教材との使用語彙で重なりを調査するため、それぞれ約 113,000 語と、約 14,000 語に及ぶコーパスを構築した。分析結果からは、各教材で約 750 ～ 2,000 種類の目標語が TOEIC 問題で出現することが明らかとなった。また、頻繁に使われる 4,000 語レベルの語彙を習得することで、当該テストで使われる語彙の少なくとも約 70% を理解できることが推計された。

キーワード

民間試験、語彙習得、コーパス分析、TOEIC Listening & Reading

Abstract

University students are urged to take Test of English for International Communication（hereinafter called TOEIC）; a larger number of universities have given such preferential treatments as recognition of credits on certain scores. Meanwhile quite a few companies require new graduates to submit their test scores as a part of the applications, lots of employees are set terms of defined scores for their advancements. However, according to the reports by IIBC（2019）, the average score of Japanese examinees is 520 in 2018; that has been lower in the ranking of Asian countries. The present study investigated used vocabulary both in TOEIC and vocabulary-building books. Then, a corpus derived from a total of 12 sample tests was constructed while another corpus contained the indices in the five wordbooks; the total number of words was respectively about 113,000 and 14,000. The result of analysis clarified that approximately 750 to 2,000 target words in the wordbooks appear in the sample tests. Moreover, this estimated figures explains that the most frequent words ranked as the top 4,000 will enable learners to comprehend about 70% of vocabulary used in the current tests.

Key words

commercial language tests, vocabulary acquisition, corpus analysis, TOEIC LR

１．はじめに

１．１　英語教育改革とその背景

文部科学省では時代の変化に伴う国際化への対応を目的として、外国語教育の見直しを求めているが、こうした行政主導の教育改革の動向は、今回に始まったことではない。臨時教育審議会による1989年の答申では「外国語で積極的にコミュニケーションを図ろうとする態度の育成」が打ち出され、新設科目「オーラル・コミュニケーション」を導入する契機となった。文部科学省（2002）からは平成14年の「『英語が使える日本人』の育成のための戦略構想」が決定し、翌年からの５カ年計画の中でSELHi（Super English Language High School）による先進的カリキュラム推進校の設置や、大学入試センター試験におけるリスニング・テストの実施が追加された。そして2020年度には、小学校高学年で必修だった「外国語活動」が中学年へ移行し、高学年は教科化される「外国語（英語）」の全面実施を迎え、現場と教育専門家を含めて様々な方面から関心を集めている。

今や中等教育では平成28年度の中学校学習指導要領の改訂に

聖徳大学文学部文学科・准教授

続き、平成29年には新しい高等学校学習指導要領が公示された。中学校では既に移行措置に伴う新課程の先行実施が始まり、高等学校も2022年（令和４年）度から学年進行で新課程へ移行する。従来まで中等教育における英語授業は、文法知識や読解力の養成に偏っていたことが指摘され、長時間の学習で相当の精力を費やしているにもかかわらず、ほとんどの学習者が実践的な英語運用力を身に付けていないことが批判されてきた（臨時教育審議会, 1986）。これからのグローバル時代を見据えると、従来の指導で求めてきた「読むこと」、「聞くこと」だけでなく、自らの立場・主張から意思伝達でき、相互理解を深める能力が必要とされ、発信力と統合した指導にも注力すべきことが提言されている。こうした外国語コミュニケーション能力の育成に向け、文部科学省では欧州評議会（Council of Europe）が示す外国語運用能力尺度「ヨーロッパ共通参照枠」（Common European Framework. 以下CEFR）の活用を推奨し、「外国語教育の抜本的強化のイメージ」を指標とすることで、新課程の成果目標を示している（図１参照）。

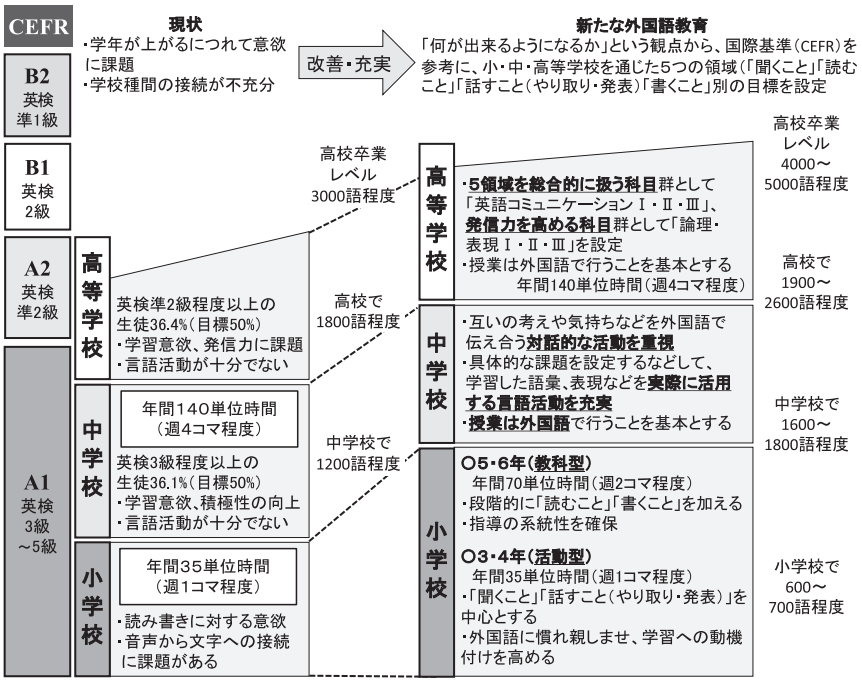
この図は「平成30年度英語教育実施状況調査」（文部科学省, 2018）の分析に基づき、対象校（3,354校）に所属していた高校３年生の英語学習到達度を示唆している。その調査結果によると、実用英語技能検定試験の取得級からCEFR Ａ２レベル（英検準２級）またはそれ以上の生徒は20.5%と報告されている。加えて、資格検定試験を未受験だがＡ２レベル相当またはそれ以上の英語力の生徒についても19.7%と推定し、双方の合算値から４割前後の生徒が一定水準の学習成果に達していることを示していた。ここで取り上げているCEFR Ａ２レベルとは、概ね「基礎的な英語知識が備わり、日常的な会話が分かり、時間をかければ短く簡単な表現でコミュニケーションできる」程度である（吉島・大橋ほか, 2004）。また、行政や経済界からの声を推察すれば、ゆくゆくは自律して外国語を学習し、海外であらゆる事態に対処し、自身の意思表明・説明といった対話のできるＢ１レベル以上の育成を要請していることが窺える。

また、当該の「強化イメージ」で着目すべきは、習得すべき語彙の大幅な増加が挙げられる。図の右端を参照すると、小学校から最大700語という新たな起点が置かれ、同時に中学校では1,600～1,800語程度（現状+600）、そして高等学校では1,800～2,500語程度（現状+700）へと引き上げられることに気づく。要するに、高校卒業までの目標数値を4,000～5,000語程度に掲げたのだが、語彙力の急激な増加を求める姿勢に対し、鳥飼（2018）は教師側が躍起となって生徒に英単語を暗記させたり、単語テストばかりを行う現場の混乱を危惧している。専門家の視点から、彼女もまた英語の運用力に語彙力が不可欠であることは認めているが、脈絡のない暗記の強要は学び続ける意欲を奪うリスクを指摘しているのである。

１．２　大学における英語教育の現状

初等・中等教育での改革が進む一方、高等教育を担う大学においては、各校が独自に掲げる教育理念や目的に基づいてカリキュラムを編成するため、一括りに英語力の到達目標を議論することが難しい。とはいえ2000年前後から叫ばれるようになった大学生の学力低下の一端は、ベネッセ教育総合研究所（2014）より報告されており、大学の入学前教育を実施していた割合は

図１　外国語教育の抜本的強化のイメージ（文部科学省, 2018）



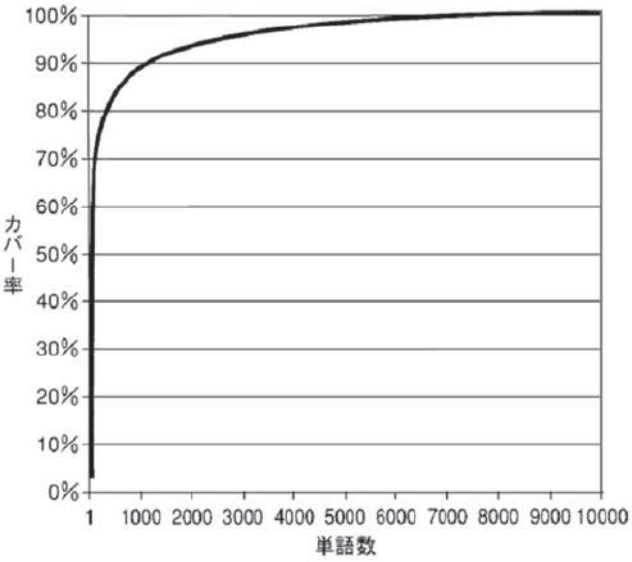
調査対象の37.1%、リメディアル教育に関しても34.4%が実施していたことが確認できる。こうした実情から、高大接続のあり方に対する問題が見直される一方で、入学後の確かな学力という意味での英語力について、大学側は学生への明確な指針を与え、課題を解決させていくべきだろう。

現在では多くの大学が入学生に民間試験の受験を奨励し、例えば高得点者に英語科目の単位を認定したり、科目成績の一部に充てる等の活用事例も報告されている（IIBC, 2016）。だが民間試験の受験者が増加する一方、世界規模で日本人一般の英語力は未発達だとする意見は多い。IIBC（2019）ではEducational Testing Service（以下ETS）が主催する2018年TOEIC Listening & Reading Testの国・地域別平均スコアがまとめられ、日本は平均520点と報告している。この値は掲載した49の国と地域で44番目に位置し、近隣のアジア諸国の中華人民共和国や台湾よりも低く、韓国に至っては同年のスコアで153点もの開きがあった。こうした実情は筆者が把握する限り10年以上続いており、高等教育機関である大学の役割を再考すべき時期にあると考える。とはいえ、英語専攻やそれに類する学部・学科を除けば、国内の大学は外国語科目を１～２年次までででカリキュラム編成する場合が多く、大抵の語学教師には現状以上の成果を求めることは難しいかもしれない。鳥飼（2018）もまた、大学で民間試験の得点を上げること自体が目的となっては、英語を学ぶ動機づけに繋がらず、コミュニケーションに使える英語の学びは望めないと指摘している。ここで重視すべきは自律性と継続性を兼ね備えた学習者の基盤づくりであり、望むべくは、学習段階の節々にも成果を実感できる指導プログラムの整備が意欲の持続にも必要だろう。

１．３　語彙習得段階の可視化

大学生の低学力傾向が顕在化し、リメディアル教育の必要性が高まった2000年代以降、基礎項目に偏った指導に終始する英語授業が多くの大学で行われてきた。酒井ほか（2010）による現場の忸怩たる思いの一節を紹介すると、大学教科書には英検４級程度のものが氾濫し、却ってそれがもてはやされていた当時の様子が綴られている。しかし彼らはその後述で、大学入学後も中学校や高校での基礎を復習させたとしても、積極的な意味を見出せないと説き、主体性を持つような学習支援の必要性を提案している。確かに多くの低学力な学生を抱える現場で陥りやすい状況として、学習者の能力や意欲を根深い特性と思い込み、平易な教材を与え続けてしまうと、負の循環を助長するだけで自律的な学習段階に進むことはできない。加えて、こうした教材は基礎文法の例文が並ぶばかりで、語彙レベルが抑えられ、内容の豊かさや味わいのある表現に乏しい。文法知識を体系的に学ぶことに異論はないが、次なる段階を目指す材料として一定量の語彙は必要なのである。しかし学生の自学自習に任せる教師や、個々の語彙力を把握せず単語集を購入させるだけの実態も多く、どのくらいの語彙を覚えれば実用的な英語に到達するのか多くの語学教師間で周知されていない。望月・相澤・投野（2003）はBritish National Corpus（BNC）による検証を紹介し、英語母語話者が実際の会話で使う語彙の傾向について説明している（図２参照）。それによると、日常会話で80%以上は最も頻繁に使われる500語であり、範囲を最頻1,000語まで広げると89%、そして2,000語を上手に使いこなせば94%を賄えると結論づけている。仮に高等学校の在学中で4,000語の習得に間に合わなくとも、最頻2,000語+αの語彙を確保し、残り2,000語に触れる機会を効率的に与えれば、実質２カ年の大学授業で内容のある言語活動を実践し、必要な語彙を取り戻すことは可能であろう。

図２　BNC話し言葉の1万語コーパスに対するカバー率（望月・相澤・投野, 2003: p. 27）



国内の言語習得に関わる研究もまた、コンピュータ処理技術の普及と、英文データの膨大な蓄積によるコーパスの構築が可能となった90年代から飛躍的に発展してきた（投野, 1997）。大学英語教育学会（JACET）教材研究委員会では、1981年と1983年の試案を経て、1993年の基本語リスト編纂に至った。その編纂にあたっては、先述のBNCを基準データとしながら、日本人の英語学習の実態に合わせて国内メディアや教材のデータを取り込み、入試問題の語彙検定の指標としても用いられている（JACET基本語改訂委員会, 2003）。現在ではJACET 8000と呼ばれるCD-ROMソフトやWeb環境での分析ツールが公開され、任意の英文をデータ投入することで、語彙レベルを瞬時に分析することが可能である。

１．４　本研究の目的

前述の民間試験の１つであるTOEIC Listening & Readingは、

国内でも受験者人口が増大し、自身の現場で関心を持つ学生も増えつつある。しかし日常会話のみならずビジネス場面を想定した問題には、習熟度だけでなく社会経験の少ない学生には頭を悩ませることも多い。こうした教育現場の実情を踏まえて留意すべきは、まず学習の目標に適った教材を提供し、必要な語彙に接触し、活用の機会を増やすことと考える。本研究では大学生が入学前後で手にするであろう学習教材を調べ、それらがどのくらいTOEICテストで用いられる語彙をカバーしているのか推定する。一連の調査を通じ、対象の教材が学習者にどのくらい多様な語彙知識を与え、またTOEICテストがどのくらい複雑で高度な語彙を要求しているのか考察する。

2．方法

2．1　対象テキスト

2．1．1　TOEIC公式問題集

本研究の主たる対象であるTOEICテストの分析には、ETSより市販されている公式問題集[1)]を用いた。同テストは2016年５月以降より変更され、「公式TOEIC Listening & Reading 問題集－新形式問題対応編」（ETS, 2016a）から現行の出題方式に対応している。今回はより安定した標本抽出を図り、同シリーズの「問題集１」及び「問題集２」（ETS, 2016b; 2017）を対象に加えた。また、同テスト内での比較と併存性を調べるため、旧方式（2006年３月～2016年３月以前）に対応の「TOEICテスト新公式問題集」（ETS, 2005）及び同じシリーズからvol. ２とvol. ３の計３冊を用いた（ETS, 2007; 2008）。なお、新旧でテスト方式の変更点は下の表１に示したとおりである。

表１　TOEIC新旧でテスト方式の変更点

	Part：タスク形式	旧方式	新方式
Listening（約45分）	Part 1：写真描写問題	計10問	計6問
	Part 2：応答問題	計30問	計25問
	Part 3：会話問題	計30問	計39問
	Part 4：説明文問題	計30問	計30問
Reading（75分）	Part 5：短文穴埋め問題	計40問	計30問
	Part 6：長文穴埋め問題	計12問	計16問
	Part 7：読解問題	計48問	計54問

表で記載された質問数の変更以外に、まずリスニング・セクションのPart ３及びPart ４で問題の一部でグラフ等の視覚情報が加わった。またリーディング・セクションに関しても、Part ７で複数の文書を読み解く問題（multiple-passage tasks）の配分が増加した点が挙げられる。

2．1．2　語彙学習教材

前述のとおり、TOEICテストへ対応した指導には、学習者の習得語彙を平素より確認・補強するための学習教材が必須である。とりわけ日本人学習者の使用を考慮に入れ、国内で市販されている単語集に絞った。表２は今回の対象書籍の一覧であり、大学進学を視野に入れた高校生から、大学生、そして社会人のTOEIC受験者が用いている主要タイトルとして選定した。選定条件として、習得目標の語彙が例文と一緒に提示され、文脈効果などで語義の理解がしやすく、記憶保持にも期待できるものとした。ただし各書籍の編集コンセプトには様々な特色があり、単一の例文に目標語句１つを提示したものがあれば（書籍Ⅰ，Ⅳ，Ⅴ）、例文毎に複数の目標語句を埋め込んだものもみられた（Ⅰ，Ⅲ）。また慣用表現や群動詞といった成句は書籍によって取扱いが異なり、例文を与えたもの（Ⅲ，Ⅴ）、注釈のみで定義したもの（Ⅰ，Ⅱ）、全く言及しないもの（Ⅳ）に分かれた。そこで、今回の分析では指標の統一を図り、巻末のインデックス（index）に掲載された語彙・成句を分析対象とした。

表２　対象教材と収録項目の一覧

書籍名（出版社）	目標語彙	成句	例文	フレーズ
Ⅰ．速読英単語①必修編（Z会）	1519	―	744	―
Ⅱ．ターゲット1900（旺文社）	1900	―	1500	400
Ⅲ．DUO3.0（アイシーピー）	1572	997	560	―
Ⅳ．LINKS1500（金星堂）	1500	―	1500	―
Ⅴ．TOEIC Listening & Reading 公式ボキャブラリーブック（ETS）	1000	155	1155	―

2．2　データ入力と分析

データの下準備として、コンピュータ上での言語情報処理に適したコーパスを作成する必要がある。そこでTOEIC公式問題集からは問題文と質問文、そしてリスニング・セクションの音声スクリプト全てをプレーン・テキスト形式（.txt）の電子ファイルとして作成・保存した。語彙学習教材のインデックスも同様、全ての語彙・成句を電子ファイルへ転記した。なお、出題前に放送または読まれる指示文や問題番号、インデックスのページ番号はコーパス・データから除外した。

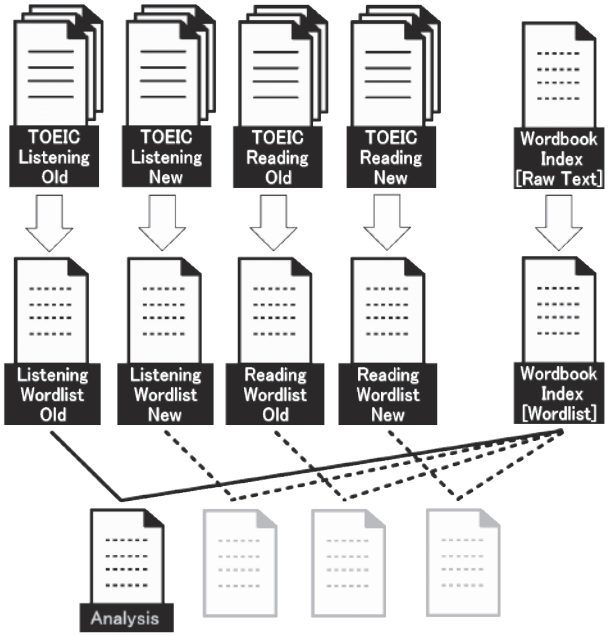
語彙の多様性と複雑さを調べるにあたり、多くの日本人学習者が習得に困難な語彙の検出を目的とし、オンライン解析ツールJACET 8000 Level Analyzer（清水, 2015）を用いた。このコーパス・ツールは、CGI機能を搭載したウェブサイトの形体を採り、所定のテキストボックスから調べたい英文を送信することで、大学英語教育学会基本語リスト（JACET 8000 List of 8000 Basic Words）に基づいた使用語彙の評価をする。集計結果には、異なり語数（types: lexical diversity）と延べ語数（tokens）が表示され、英文に含まれた語彙の複雑さ（lexical complexity）は９段階で表示される。JACET 8000の語彙目録

では、最も高頻度で出現する基本語の1000語をLevel 1、最高レベルとする低頻度語をLevel 8で表し、語彙目録に載っていない単語はOver 8として別途集計される。

2.3　手順と出力データの解釈

電子データ化された対象テキストは、第一段階としてTOEIC公式問題集の新版と旧版に分けて投入し、同様に単語集のインデックスを個別に投入する。それぞれの解析データには、重複して出現した単語を1つにまとめたワード・リストが出力されるため、第二段階では公式問題集と単語集インデックスを基にしたワード・リストを同時投入する。一連のデータ投入から出現頻度2を示した単語は、当該のTOEIC問題と単語集に重複して出現した語彙として絞り込むことができる（手順のイメージは図3を参照）。本研究の考察では、これら双方のコーパス・データより重複出現した語彙の数を、単語集から学習できる語彙の多様さ・複雑さの指標として扱う。

図3　コーパス分析手順のイメージ



3.　結果と考察

3.1　対象テキストの基礎計測

3.1.1　TOEIC公式問題集

表3は対象テキストであるTOEIC公式問題集より、リスニング・セクション（Part 1～4）の旧版（Old）と新版（New）の各3冊を抜粋し、語彙解析した結果の要約である。表の左半分には異なり語数（Types）の計測値と語彙レベルごとの割合（%）、右半分には延べ語数（Tokens）による計測値をまとめ、新版から旧版で差し引いた値（N-O）を最右列に記し、双方でどのくらい語数に増減があるのか推計した。

そこで異なり語数から新旧で比較した場合、概ね新版で増加こそみられたが、全体でも129 typesの変化に留まり、各レベル別の差し引きが多い箇所でも34 types（Lv 5）と微々たる変化だった。また、延べ語数では旧版の25,267 wordsに対し、新版が27,267 wordsを数え、模擬テスト各6回分の累積で2,000 wordsの増加がみられた。この値からテスト1回分に換算すると約330 words程度の違いであるが、新方式のリスニング問題に移行して若干の増量が推定される。しかしながら、レベル別に語彙のカバー率を概観すると、計測値のほとんどが1%に満たない増減に留まっていた。数少ない1%超の変化としては、Type Lv 5（+1.33%）にみられる新版での低頻度語彙の増加、そしてType Lv 1（-2.28%）とToken Lv 1（-1.94%）にみられる基礎語彙の微減が確認された。そこで語彙レベル別の数列からピアソン積率相関係数を算出すると、表中の最上段に示したとおり極めて1.00に近似した相関[2)]が得られた。ここまでの分析結果よりTOEICリスニング・セクションについて述べるなら、問題作成で極めて統制された語彙使用を施し、項目困難度の等質化を図っていることが窺える。

表3　リスニング・セクションにおける語彙レベル間の増減と相関

	Types(r = 0.998)					Tokens(r ≒ 1.000)				
	Old	(%)	New	(%)	N-O	Old	(%)	New	(%)	N-O
Lv 1	792	(38.19)	791	(35.91)	-1	21319	(84.37)	22476	(82.43)	1157
Lv 2	500	(24.11)	517	(23.47)	17	2214	(8.76)	2578	(9.45)	364
Lv 3	294	(14.18)	322	(14.62)	28	806	(3.19)	1011	(3.71)	205
Lv 4	167	(8.05)	178	(8.08)	11	373	(1.48)	470	(1.72)	97
Lv 5	76	(3.66)	110	(4.99)	34	144	(0.57)	255	(0.94)	111
Lv 6	48	(2.31)	65	(2.95)	17	86	(0.34)	153	(0.56)	67
Lv 7	28	(1.35)	28	(1.27)	0	91	(0.36)	53	(0.19)	-38
Lv 8	24	(1.16)	16	(0.73)	-8	35	(0.14)	34	(0.12)	-1
Over 8	145	(6.99)	176	(7.99)	31	199	(0.79)	237	(0.87)	38
Total	2074		2203		129	25267		27267		2000

表4　リーディング・セクションにおける語彙レベル間の増減と相関

	Types(r = 0.995)					Tokens(r = 0.995)				
	Old	(%)	New	(%)	N-O	Old	(%)	New	(%)	N-O
Lv 1	832	(26.62)	830	(26.26)	-2	21171	(71.94)	23065	(72.40)	1894
Lv 2	625	(20.00)	635	(20.09)	10	4050	(13.76)	4282	(13.44)	232
Lv 3	528	(16.90)	493	(15.60)	-35	1903	(6.47)	1988	(6.24)	85
Lv 4	289	(9.25)	282	(8.92)	-7	761	(2.59)	793	(2.49)	32
Lv 5	217	(6.94)	204	(6.45)	-13	449	(1.53)	405	(1.27)	-44
Lv 6	118	(3.78)	127	(4.02)	9	240	(0.82)	264	(0.83)	24
Lv 7	69	(2.21)	69	(2.18)	0	216	(0.73)	258	(0.81)	42
Lv 8	47	(1.50)	54	(1.71)	7	86	(0.29)	68	(0.21)	-18
Over 8	400	(12.80)	467	(14.77)	67	552	(1.88)	733	(2.30)	181
Total	3125		3161		36	29428		31856		2428

表4ではリーディング・セクション（Part 5～7）旧版（Old）と新版（New）の分析結果を示す。前掲の表3と同様、左半

分には異なり語数、右半分には延べ語数による計測値と割合（%）をまとめ、双方の語数の違いを各最右列（N-O）に表示した。異なり語数に基づいて比較すると、新版の方が若干種類こそ多かったが、全体でも36 typesの違いにすぎなかった。延べ語数の比較からは旧版の29,428 wordsに対し、新版は31,856 wordsを数え、サンプル・テスト各６回分の累積で2,428 wordsの増加が確認された。テスト１回分に換算すると約400 words程度の違いであるが、これは新方式のPart ５の短文穴埋め問題が減少し、Part ７で提示される長文の数が増えたことで容易に推測される。レベル別に語彙のカバー率を概観すると、どの計測値も１%前後の増減に留まり、当該の数列同士でピアソン積率相関係数を求めると、リスニング・セクションと同様、1.00に極めて近い相関だった。一連の検証から、TOEICは約10年を経て出題方式の変更を公表したが、使用語彙については問題作成の統制が取られ、極めて安定した項目困難度の維持に努めていることが推察された。

３．１．２　語彙学習教材

表５は語彙学習教材のインデックスをコーパス解析した結果の要約である。掲載されている語彙と成句に関して、JACET 8000のレベルごとに分類し、その異なり語数を上段、そして割合（%）を下段に表示した。各レベル（Lv １～８）の最大値は1,000 typesで、それに近い値ほど当該レベルの語彙をカバーした教材と解釈できる。掲載項目の合計（Total）から全体を見渡すと、中段の書籍Ⅲが5,399 typesと最も多く、下段の書籍Ⅳ・Ⅴが1,500 typesを下回る少なさだった。表中の下線で示すとおり、特に計測値の高い語彙レベルを検討すると、書籍Ⅰ・Ⅱ・ⅢはLevel ２からLevel ５までの語彙で掲載数が少なくとも300を上回り、大学入試問題の英語で頻出とされる5,000語の習得に重点が置かれていることが窺える。当該の３冊は全て大学入学試験の準備を視野に入れた参考書であり、書籍の序文には国公立大学の二次試験や私立大学の最難関校受験への対応が謳われている。特筆すべきは、書籍ⅢがJACET 8000のカバーしない極めて複雑な語彙1,000語超を掲載し、Level ８の229 typesを除けば全てのレベルで３割以上を網羅している点にある。一方、書籍Ⅳは重点項目として2,000～3,000語レベルを半分近くカバーしているが、Level ５以降の低頻度語は１割にも満たないカバー率であった（Types < 100）。この書籍は前者３タイトルと異色の編纂プロセスを経ており、大学で使われている英語教科書からの語彙が主に選定されている。参考文献の非公表により引用元が明らかではないが、往々にして日本人学生向けに作られた教科書は、抜粋した専門書からの用語の平易化（simplification）が施されることも多く、低頻度かつ複雑な語彙は選定リストに挙がらなかった可能性が考えられる。また、書籍Ⅴは掲載項目が最も少なく、Level １～３の基礎語彙から英字新聞や雑誌などの一般メディアで頻繁に使われる語彙で約３割、Level ５以降の専門性の高い語彙のカバー率は１割未満だった。この書籍はTOEICテストに合わせた公式教材のため、筆者自身が予見していた結果とは異なる低い値だった。

表５　教材インデックスに基づく語彙レベルの分布

	Lv 1	Lv 2	Lv 3	Lv 4	Lv 5	Lv 6	Lv 7	Lv 8	Over 8	Total
Ⅰ	255	509	564	399	322	169	166	104	391	2879
(%)	(8.86)	(17.68)	(19.59)	(13.86)	(11.18)	(5.87)	(5.77)	(3.61)	(13.58)	(100.00)
Ⅱ	119	444	588	526	480	292	246	150	518	3363
(%)	(3.54)	(13.20)	(17.48)	(15.64)	(14.27)	(8.68)	(7.31)	(4.46)	(15.40)	(100.00)
Ⅲ	732	777	794	619	543	366	322	229	1017	5399
(%)	(13.56)	(14.39)	(14.71)	(11.47)	(10.06)	(6.78)	(5.96)	(4.24)	(18.84)	(100.00)
Ⅳ	158	456	437	209	113	50	17	13	43	1496
(%)	(10.56)	(30.48)	(29.21)	(13.97)	(7.55)	(3.34)	(1.14)	(0.87)	(2.87)	(100.00)
Ⅴ	322	312	309	161	85	39	18	9	46	1301
(%)	(24.75)	(23.98)	(23.75)	(12.38)	(6.53)	(3.00)	(1.38)	(0.69)	(3.54)	(100.00)

３．２　収録語彙の照合と比較

ではTOEIC対策で語彙力の補強を試みた場合、どの教材から語彙習得の有効性が期待されるだろうか。表６と表７は公式問題集と教材インデックスによるワード・リストを解析ツールに同時投入し、各セクションと教材との組合せで重複した語彙の頻度をまとめた一覧である。まずリスニング・セクションに対する教材インデックスで語彙の重なりを検討すると、各教材との頻度がLevel ２を中心に多くみられ、その周辺から徐々に減少していることに気づく。また書籍Ⅲ・Ⅴでは数多くの成句もインデックスへ掲載されていたが、それがLevel １で高い頻度を示していたと考えられる。つまり、それは慣用表現である句動詞や副詞句といった学習項目が関係しており、基本語彙で構成しながら理解・習得に難しいフレーズ[3)]に配慮していた結果といえるだろう。最下段の値から全体の語彙重なり（Total）を比較した場合、書籍ⅢがTOEICテストにおける出現語彙を最も多く網羅し、大学入試対策に特化した書籍Ⅰ・Ⅱは前者の半分程度、そして書籍Ⅳでカバーできる語彙は最も少なかった。

表６　語彙重なり①：TOEICリスニング × 教材インデックス

	Ⅰ. 速読英単語 ①必修編		Ⅱ. ターゲット 1900		Ⅲ. DUO 3.0		Ⅳ. LINKS 1500		Ⅴ. TOEIC公式 ボキャブラリー	
	Old	New	Old	New	Old	New	Old	New	Old	New
Lv 1	166	169	88	92	601	598	109	111	302	308
Lv 2	267	285	226	252	402	415	221	236	226	238
Lv 3	164	177	166	200	248	266	125	146	146	178
Lv 4	51	50	82	83	94	96	32	37	75	76
Lv 5	22	21	26	39	39	51	10	11	31	41
Lv 6	8	4	14	10	13	18	3	5	16	24
Lv 7	2	2	1	4	4	6	2	2	3	6
Lv 8	3	0	3	2	6	1	0	0	2	3
Over 8	1	1	6	6	15	17	1	2	14	15
Total	684	709	612	688	1422	1468	503	550	815	889

表７　語彙重なり②：TOEICリーディング × 教材インデックス

	Ⅰ. 速読英単語①必修編		Ⅱ. ターゲット1900		Ⅲ. DUO 3.0		Ⅳ. LINKS 1500		Ⅴ. TOEIC公式ボキャブラリー	
	Old	New	Old	New	Old	New	Old	New	Old	New
Lv 1	193	179	104	92	**633**	**625**	129	120	**315**	**314**
Lv 2	**377**	**369**	**329**	**327**	**531**	**531**	**316**	**315**	285	281
Lv 3	**320**	284	**331**	**301**	**430**	**406**	245	227	260	252
Lv 4	131	115	179	153	195	174	70	57	115	120
Lv 5	74	61	97	85	123	105	18	15	51	54
Lv 6	22	19	35	40	40	44	12	12	27	28
Lv 7	13	10	11	16	25	22	3	1	8	12
Lv 8	4	9	8	8	8	15	0	2	3	5
Over 8	17	7	16	13	39	32	3	1	16	19
Total	**1151**	**1053**	**1110**	**1035**	**2024**	**1954**	796	750	**1080**	**1085**

注．Ⅰ＝速読英単語①必修編，Ⅱ＝ターゲット1900，Ⅲ＝DUO3.0，Ⅳ＝LINKS1500，Ⅴ＝TOEIC Listening & Reading 公式ボキャブラリーブック，太字・下線は特に計測値の高い語彙群（Listening: Lv > 200, Total > 800; Reading: Lv > 300, Total > 1,000）.

何よりも特筆すべきは、書籍Ⅴの学習項目数が対象の５冊の中で最少にもかかわらず、全レベルを合わせた語彙重なりが書籍Ⅲに次いで高い値だったことが挙げられる。

既に表３と表４で延べ語数が示したとおり、TOEICリスニング及びリーディング・セクションで比べた場合、コーパス・サイズで旧版は約4,100 words（Old R-L: 29,428－25,267＝4,161）、新版は4,600 words（New R-L: 31,856－27,267＝4,589）もの違いがあった。言い換えれば、リーディング・セクションで英語に触れる量は多く、その試験時間も相応に長く設定されている。また異なり語数の比率を算出した場合、リーディング・セクションで用いられる語彙の種類はリスニングの約1.4～1.5倍に達することが分かった（Old R/L: 3,125／2,074 = 1.55, New R/L: 3,161／2,203 = 1.43）。そこで表７のリーディング・セクションに対する教材インデックスとの重なりに関しては、高頻度とする水準を語彙レベルで300以上、全体数で1,000以上に引き上げて検討した。これに基づいて頻度の高い語彙レベルを見渡すと、概ねリスニング・セクションに似た分布を描き、Level ２を軸とした重なりが書籍Ⅰ～Ⅳから確認された。また書籍Ⅳを除けば、いずれの教材もLevel ２からLevel ４にかけて100以上の重複を数え、これらの教材を通じて習得された語彙はリーディング・セクションの方でむしろ遭遇する確率が高いとみられた。最下段から全体の語彙重なりを調べると、書籍Ⅳを除く４冊が1,000 words以上をカバーし、書籍Ⅲに至っては他の２倍近くを網羅することが推計された。また、TOEICテストの公式教材である書籍Ⅴは、インデックスの収録語1,311 wordsに対し、約82%が公式問題集と重複していた（1,080 / 1,311 = .824）。当然の結果ながら、この書籍がいかに当該テストに焦点を合わせた語句選定であったか実証したといえる。

４．まとめ

４．１　TOEICテストの求める語彙力とは

本研究ではコーパス・ツールを用い、TOEICテストで求められている習得語彙の推計に取り組んだ。既に３．１の分析で触れたとおり、TOEIC Listening & Reading Testは2016年以前の方式と比較して、英語の言語接触量が増加傾向であることが分かった。その一方、使用語彙の多様性や複雑さを検討すると、レベル間の配分・比率が極めて安定し、これがTOEICテストの実施回を超えたスコア等質化に関係していると考えられる。具体的な語彙レベル指標としてJACET 8000で検討した場合、中学校卒業までに学ぶ約1,000語の語彙基盤を固め、次いで高い頻度で用いられる2,000～4,000語レベルを習得していれば、少なく見積もってもリーディング・セクションの異なり語数（types）で約70%[4)]をカバーすることが分かった。この4,000語という語彙習得はCEFRのB１レベルに相当し、新課程における高等学校卒業までに望まれる習熟度として掲げられた目標とも結びつき、当該試験の受験準備が可能な目安と位置付けられるだろう。

なおCEFR B１レベルを他の資格検定試験の到達度で照らし合わせた場合、実用英語技能検定で算出されるCSE（Common Scale for English）スコアで約2,000点前後とされ、２級合格者の上位得点者から準１級への受験を視野に入れた到達度までに相当する。必ずしも語彙力が英語運用力の全てを決定づけるわけではないが、より多くを理解できる語彙知識を身につけることは、読解の精度を高めたり（Laufer, 1992）、未知の単語に遭遇した際の推測にも役立つ（Huckin & Coady, 1999）。こうした自律的な学習者を育成するためには、語学教師の立場から個々の学生の到達度に適した方法を助言できる専門理論や経験的知見を蓄積したいところである。

４．２　語彙学習教材の選定と活用について

今回の教材による分析結果を振り返ると、TOEICテストに狙いを絞った単語集に限らず、従来より市販されている大学入試参考書の類に関しても、かなりの割合で語彙力の補強に期待できることが分かった。特にTOEICテストの語彙の複雑さは2,000語と3,000語レベルでの比重が高く、双方で用いられている語彙との重なりが多い箇所では500を超えるケースさえ確認された（例. Lv ２：リーディング×書籍Ⅲ）。その一方で、複雑かつ高度な語彙はTOEICテストも例外ではなく出現頻度が低く、対象教材との重なりも高レベルほど減少することは避けられない。しかし、語彙学習教材の限界を悲観的に捉えるべきではなく、先述４．１で触れたとおり、獲得した多様な語彙知識ほど推測して理解できる能力に導くのであれば、実践的な聴解・読解と併用し、発展した学びに取り組ませるべきだろう。

また、高度で複雑な語彙だけでなく、2,000語レベル以下の

基本語で構成される多様な成句の提示に注力した教材が有用であることも再認識すべき結果だった。とりわけ英語の成句とは、日本の受験英語に特有の「英熟語」もしくは「イディオム」と称し習得の難しさが指摘され（副島, 1997）、文法や語彙とは別個に取り扱った参考書が数多く出版されてきた。これは日本語・英語の語彙が一対を成した逐語訳で理解し続けてきた初期の学習時期から、より高次の段階に進む学習者にとって困難に陥りやすい課題でもある。そうした語彙習得の複雑さや課題を明示するか否かの編集方針で、今回の対象教材は別側面からの分岐点を顕わしていたともいえるだろう。

4.3　その他の教育的示唆

本研究で分析対象とした教材は、いずれも音声教材が別売または無償のインターネット配信で用意されており、とりわけTOEICリスニング・セクションでは学習成果を発揮する上で必須であることに疑う余地はない。往々にして、初歩の英語学習者ほど必要最小限の情報に絞りがちで、綴りや主な語義のみを記憶することに終始することが多い。しかし、文字と音声の一体化に限らず、目標語彙の派生語、類義語や互換可能な成句にも余すことなく掲載スペースを充てた書籍があったように、ことばとは表現の豊かさ故に、簡単ではない語彙習得の本質から避けてはならないだろう。

最後に語学教師が留意すべきこととして、学習者の習熟度を把握できる尺度を持ち、意識的に学習者自らに目標を再検討させることも重要である。手軽で安価な例を挙げるなら、手元の単語集から成果を引き出せる英字新聞やメディアを図書館等で利用させたり、大学や地域の外国語コミュニケーション環境に触れさせたりすることを勧める。また、本研究の主たる対象であったように、習熟段階に合った民間試験に臨むことで活用度を増すことも現実的な手段といえるだろう。いずれの方法を採るにせよ、語学教材や書籍を手に取る際は、現状の習熟段階と照らし合わせ、目標に適ったものを選択することが有意味かつ持続できる語学への第一歩と考える。

注.

1）TOEIC運営委員会によると、異なる実施回の間で統計的なスコア同一化（equating）の処理をするために過去問題を非公表としている。したがって、一般的なサンプルの入手方法はETS公式問題集が通常とされる。

2）相関係数の有意確率は、全て0.01%水準をはるかに下回る値だった（Listening types: $p = 6.13 \times 10^{-10}$, tokens: $p = 1.03 \times 10^{-14}$, Reading types: $p = 3.15 \times 10^{-08}$, tokens: $p = 3.44 \times 10^{-15}$）。

3）Level 1の基本語だけを含む成句として、例えば書籍Ⅲにはaround the clock, fall short of, have second thought(s)、書籍Ⅴにはin person, out of order, report to work等が掲載されている。

4）延べ語数（tokens）で換算した場合、語彙レベルの高いリーディング・セクションの推計でも約94%をカバーしていた。

5．引用文献

Educational Testing Service（2005）「TOEICテスト新公式問題集」国際ビジネスコミュニケーション協会

ETS（2007）「TOEICテスト新公式問題集２」国際ビジネスコミュニケーション協会

ETS（2008）「TOEICテスト新公式問題集３」国際ビジネスコミュニケーション協会

ETS（2016a）「TOEICテスト公式問題集 新形式問題対応編」国際ビジネスコミュニケーション協会

ETS（2016b）「公式 TOEIC Listening & Reading 問題集vol. 1」国際ビジネスコミュニケーション協会

ETS（2017）「公式 TOEIC Listening & Reading 問題集vol. 2」国際ビジネスコミュニケーション協会

Huckin, T., & Coady, J.（1999）Incidental vocabulary acquisition in a second language. A Review. Studies in Second Language Acquisition, 21, 181-193.

IIBC（2016）TOEIC Tests 入学試験・単位認定における活用状況－大学・短期大学・高等専門学校－ [Website] Retrieved from https://www.iibc-global.org/toeic/official_data/lr/search.html

IIBC（2019）2018 Report on Test Takers Worldwide: TOEIC Listening and Reading Test [PDF] Retrieved from https://www.iibc-global.org/library/default/toeic/official_data/pdf/Worldwide2018.pdf

Laufer, B.（1992）How Much Lexis is Necessary for Reading Comprehension? In: Arnaud P.J.L., Béjoint H.（eds）Vocabulary and applied linguistics（pp.126-132）. London: Palgrave Macmillan.

酒井志延・中西千春・久村研・清田洋一・山内真理・間中和歌江・合田美子・河内山晶子・森永弘司・浅野享三・城一道子（2010）「大学生の英語学習の意識格差についての研究」『リメディアル教育研究』5, 9-20.

清水伸一（2015）JACET 8000 Level Analyzer [Website] Retrieved from http://mochvocab.sakura.ne.jp/cgi-bin/j８web/j８web.cgi

副島隆彦（1997）『続・英文法の謎を解く』筑摩書房

大学英語教育学会基本語改訂委員会（編）（2003）『大学英語教育学会基本語リスト（JACET8000）』大学英語教育学会

鳥飼玖美子（2018）『英語教育の危機』東京：筑摩書房

ベネッセ教育総合研究所（2014）高大接続に関する調査 [PDF] Retrieved from https://berd.benesse.jp/up_images/research/2014_koudai_all.pdf

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

文部科学省（2002）「英語が使える日本人」の育成のための戦略構想の策定について [Website] Retrieved from http://www.mext.go.jp/b_menu/shingi/chousa/shotou/020/sesaku/020702.htm

文部科学省（2019）平成30年度「英語教育実施状況調査」の結果について [Website] Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1415042.htm

吉島茂・大橋理枝（訳・編）・奥聡一郎・松山明子・竹内京子（共訳）（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

臨時教育審議会（1986）教育改革に関する第二次答申（抄）[Website] Retrieved from http://www.mext.go.jp/a_menu/shotou/senkou/1256677.htm