

令和4年度の新課程英語教科書コースおよび共通テストコースに基づく TOEIC®

Listening and Reading Test / TOEIC Bridge® Listening and Reading Tests の語彙の特徴

卯城 祐司（筑波大学）

小室 竜也（筑波大学大学院・日本学術振興会特別研究員）

目次	
1. はじめに.....	3
2. 先行研究 .....	3
2.1 言語理解に必要な語彙力.....	3
2.2 TOEIC L&R および TOEIC Bridge L&R と語彙力の関係.....	4
2.3 本研究の目的.....	4
3. 教科書コーパスに基づく TOEIC L&R、TOEIC Bridge L&R、共通テストの語彙のカバー率 .....	5
3.1 検定教科書、TOEIC L&R、TOEIC Bridge L&R、共通テストのコーパス化 .....	5
3.2 語彙の多様性の観点からのコーパス情報の解釈.....	9
3.3 レマ換算のカバー率.....	10
3.4 中高英語教科書、TOEIC L&R、TOEIC Bridge L&R、共通テストで使用されている語彙の概観 .....	12
3.5 単語・Ngram でコーパスの分類.....	14
3.5.1 ランダムフォレストによる単語の出現頻度での分類 .....	14
3.5.2 コレスポネンス分析による Ngram の出現頻度での分類 .....	16
3.6 中高英語教科書および共通テストから見た TOEIC L&R に特有な語彙.....	17
4. 結論 .....	25
引用文献.....	27
付録 1 TOEIC L&R に登場し教科書に登場しない語彙 .....	28
付録 2 TOEIC Bridge L&R に登場し教科書に登場しない語彙.....	33
付録 3 中高教科書、TOEIC L&R、TOEIC Bridge L&R、共通テストで用いられている高頻度の Ngram.....	34
付録 4 中高教科書で使用されている語彙が TOEIC L&R、TOEIC Bridge L&R、共通テストで 使用されている語彙をカバーする割合.....	34
付録 5 TOEIC L&R、TOEIC Bridge L&R、共通テストで使用されている語彙が中高教科書で 使用されている語彙をカバーする割合.....	35
付録 6 個別教科書会社ごとのカバー率一覧.....	36
付録 7 TOEIC L&R および TOEIC Bridge L&R から見た共通テストのカバー率 .....	39

## 1. はじめに

平成 29 年 3 月には中学校学習指導要領が、平成 30 年 3 月には高等学校学習指導要領が改訂された。中学校の新学習指導要領は令和 3 年度から全面的に実施され、高等学校は令和 4 年度から年次進行で実施されている。令和 5 年 3 月現在、高校 1 年生は新学習指導要領に基づいた検定教科書に基づき授業を受けている。令和 6 年には高校 2 年生が、令和 7 年には高校 3 年生がそれぞれ、新学習指導要領の下で作成された教科書で英語を学習することとなっている。これらの教科書で学習した者の中で、大学への入学を志願している者は大学入学共通テスト（以下、共通テスト）を受験することとなる。共通テストは大学が共同して同一の期日に同一の試験問題により実施される標準テストである。標準テストで重要となることは、これまでに学習した内容と連関があることである。

日本で主要な英語の資格・検定試験の一つである TOEIC® Listening and Reading Test (以下 TOEIC L&R) は、この標準テストの一つであり、受験者は年間で 200 万人を超える。TOEIC L&R は英語によるコミュニケーション能力を幅広く測定することを目的に、身近なシーンからビジネスまで、幅広い場面を対象としている。さらに項目応答理論に基づき、同一尺度上で初級から上級までの熟達度を測定することが可能である。TOEIC Bridge® Listening & Reading Tests (以下 TOEIC Bridge L&R) も同じく標準テストであり、特に英語学習初級者から中級者を対象に、日常生活における生きたコミュニケーションに必要となる「英語で聞く・読む能力」を測定する。これらの標準テストは共通テストと同様に、これまでに学習した内容に立脚していることが求められる。日本の英語学習者の多くが中高英語教科書で英語を学習した経験があることから、本調査では中高英語教科書で身に付く英語力を構成する概念の一部として、語彙に焦点を当てる。本研究は TOEIC L&R および TOEIC Bridge L&R で使用されている語彙が新課程の中高英語教科書や共通テストで使用されている語彙とどの程度対応しているか、また TOEIC L&R および TOEIC Bridge L&R に特徴的な語彙を明らかにすることを目指す。

なお、本調査では新課程の検定教科書として中学校 1 年から 3 年および高等学校 1 年の教科書を用い、高等学校 2 年および 3 年は旧課程の検定教科書であることに留意されたい。なお、累計の語彙カバー率の計算には実際の学習者が複数の教科書を使用する状況が想定しにくいと、平均値に基づいていることにも留意されたい。実際に使用したコーパスに関して、詳細は 3.1 節を参照のこと。

## 2. 先行研究

### 2.1 言語理解に必要な語彙力

新学習指導要領では高校卒業時まで、小学校からの合計で約 4,000 から 5,000 語の語彙を指導することが示されている。この語彙知識と言語技能の関係性について、Zhang and Zhang (2022) のメタ分析によると、受容語彙知識は外国語でのリーディングにおける理解と  $.57 (p < .01)$ 、リスニングにおける理解と  $.56 (p < .01)$  の相関があることが示されており、言語理解において重要な役割を果たしている。具体的には、読解テキストにおける既知語の割合が 80%以下であると全く内容を理解できず、既知語の割合が 95%から 98%以上であれば十分に理解可能であると提唱されている (e.g., Nation, 2013; Schmitt et al., 2011)。TOEIC L&R および TOEIC Bridge L&R はリスニングおよびリーディング能力を測定するため、高得点のために語彙知識は非常に重要であると言える。

語彙知識は頻度と深く関係している。the や of のように、よく見聞きする高頻度語彙は簡単な語彙であり、低頻度の難しい語彙知識を有していることは熟達した学習者であると判断できる。この考えに基づき、頻度に基づく診断テストが多く開発されてきた (e.g., VST-NJ8: Hamada et al., 2021; Vocabulary Levels Test: Webb et al., 2017)。日本人英語学習者の語彙知識を診断するテストとして開発された VST-NJ8 は、日本人英語学習者が知っておくべき 8,000 語の語彙をリスト化した New JACET 8,000 Basic Word List (Committee of Revising the JACET Basic Words, 2016) に基づき、1,000 語レベルの頻度バンドごとにランダムに抽出された語彙の意味を選択肢の中から選ぶ形式である。語彙頻度表を用いることにより語彙の難易度を客観的に把握することができ、語彙サイズテストで 4,000 語の語彙知識を有する学習者には 6,000 語レベルの語彙が多く使われているテキストは難易度が高そうであるなどの意思決定を下すことができる。

## 2.2 TOEIC L&R および TOEIC Bridge L&R と語彙力の関係

語彙知識とリーディング・リスニングの関係は有意な中程度以上の相関が見られることがメタ分析で示されており (Zhang & Zhang, 2022)、TOEIC L&R のスコアとも深く関係している (e.g., Kaneko, 2017; 水本, 2006)。Kaneko (2017) は TOEIC テスト新公式問題集 Vol. 1 から 6 を Nation (2012) の BNC/COCA ワードファミリーリストで分析し、98%のカバー率のためにはリスニングで約 4,000 から 5,000 ワードファミリー、リーディングでは 5,000 から 6,000 ワードファミリーが必要であることを示した。ワードファミリーとは、辞書の見出し語 (レマ) に派生形態素を加えたものである。例えば、happy のレマは屈折形態素を加えた happier と happiest であり、unhappy や unhappiness のような派生形態素を加えたものはワードファミリーと呼ばれる。なお、ワードファミリーとレマの換算の目安として、5,000 レマが 3,000 ワードファミリーとしてみなされる場合がある。Kaneko (2017) の結果は日本人英語学習者にとって TOEIC L&R は困難であるような印象を与える。しかし、98%のカバー率という数字は読み飛ばしても問題ない非常に低頻度な専門用語や錯乱肢も含めていることから、実際に必要となる学習者の語彙サイズはもっと低いと推定される。

TOEIC L&R と学習者の具体的な語彙サイズの関係を検討した研究は数が少なく、未解明な点は多い (水本, 2006; 望月他, 2011)。卯城・濱田 (2017) の調査 2 では望月語彙サイズテスト (相澤・望月, 2010) を用いて日本人大学生 580 名の語彙サイズを測定し、TOEIC Bridge L&R (ETS, 2007) のスコアとの関係を調査した。構造方程式モデリングの結果、以下の 3 つの予測回帰式が得られた。

- (1) 総合点推定スコア =  $0.019$  (95%CI[0.017, 0.022]) × 語彙サイズ +  $57.609$  (95%CI[48.922, 66.296])
- (2) 読解推定スコア =  $0.013$  (95%CI[0.011, 0.014]) × 語彙サイズ +  $17.079$  (95%CI[13.824, 20.334])
- (3) 聴解推定スコア =  $0.005$  (95%CI[0.004, 0.007]) × 語彙サイズ +  $43.324$  (95%CI[38.901, 47.748])

上記の回帰式に基づくと、例えば、語彙サイズが 3,000 語であった場合には TOEIC Bridge L&R の総合スコアが 115 点であると推定される。なお、TOEIC Bridge L&R は 2019 年にトータルスコアが 30 点から 100 点に変更されており、上記の予測回帰式をそのまま当てはめることはできない点に注意が必要である。

上記の先行研究は TOEIC L&R 公式問題集の語彙レベルの推定や学習者の現時点で有している語彙サイズを元にしたスコアの推定が中心であった。しかし、英語学習者の語彙サイズは動的であることを踏まえると、中学校や高校の卒業までどの程度の語彙知識がカバーできるようになるかを検討することは、教師の教育実践にとって意義深い。中條・Genung (2005) は平成 10 年度版学習指導要領下で作成された中高英語教科書と TOEIC L&R の語彙の特徴を調査し、TOEIC L&R で用いられている語彙の約 75%が教科書の語彙でカバーされることを示した。同様に、卯城・濱田 (2017) の調査 1 では平成 25 年度版学習指導要領下で作成された中高英語教科書と TOEIC L&R の語彙の特徴を調査したところ、コーパス化された全中学校検定教科書で約 47%、中高検定教科書全体で約 79%のカバー率が得られている。この結果は、教科書が TOEIC L&R の対策のためだけに作成されたわけではないことを踏まえると非常に高い数値であると言える。

## 2.3 本研究の目的

語彙知識は TOEIC L&R で重要な役割を果たすことが示されているが、これまでに先行研究で得られている結果は旧課程の学習指導要領に基づいた検定教科書を用いている。また、標準化テストとしての TOEIC L&R および TOEIC Bridge L&R の語彙的カバー率が示されたとしても、別の標準化テストと比較することなしに難易度を推定することはできない。これらの限界点を踏まえ、本研究では新課程の中高英語教科書で使用されている単語が、TOEIC L&R、TOEIC Bridge L&R、共通テストで出題される単語をどの程度カバーしているのかを明らかにすることを目指す。さらに、中條・Genung (2005) および卯城・濱田 (2017) に基づき、中高英語教科書とは異なる出現分布を持つ TOEIC L&R および TOEIC Bridge L&R に特有の語彙を抽出する。なお、中條・Genung (2005) では British National Corpus に基づく頻度表、卯城・濱田 (2017) は JACET 8,000 (JACET, 2003) をそれぞれ用いているが、これらは 2023 年現在の新課程の学習指導要領に基づく検定教科書を分析する上では不適切である。そこで本研究では、New JACET 8,000 Basic Word List を用いることとする。

### 3. 教科書コーパスに基づく TOEIC L&R、TOEIC Bridge L&R、共通テストの語彙のカバー率

#### 3.1 検定教科書、TOEIC L&R、TOEIC Bridge L&R、共通テストのコーパス化

本研究では新学習指導要領に基づく検定教科書（中学校 4 社、高校 2 社）、公式 TOEIC® Listening & Reading 問題集 Vol. 1 から 8、TOEIC Bridge L&R の各種公式問題集、令和 3 年から 5 年の共通テストの 4 種類のテキストをコーパス化した。検定教科書の選定については、中学校は採択率の高い 4 社を、高等学校は進学校において採択率の高い 2 社を対象とした。コーパス化に際しては Python 3 (ver. 3.10.2) を用いて、ファイルに含まれている日本語や記号などを削除した。得られたテキストファイルは New Word Level Checker<sup>1</sup>によって、New JACET 8,000 Basic Word List に照らし合わされた。なお、卯城・濱田 (2017) では 2003 年段階の前バージョンである JACET 8,000 Basic Word List を用いており、使用した分析アプリケーションも v8an であることに留意されたい。

New JACET 8,000 Basic Word List に基づく中学校および高等学校検定教科書コーパスの詳細情報は表 1、TOEIC L&R および TOEIC Bridge L&R、共通テストの 3 種のコーパス情報は表 2 の通りである。New JACET 8,000 Basic Word List は日本人英語学習者向けに開発されたものであり、全部で 8 つのレベルに語彙の難易度を分けている。例えば、表 1 の中学校 1 年生教科書においてはレベル 1 を表す L1 (冠詞や前置詞などの高頻度語) は全体の 72.76%を占めている。このレベルが高くなるにつれて難易度の高い語彙が収録されていると言える。表 2 に基づくと、TOEIC L&R で使用されている語彙は累積で 93.39%が、New JACET 8,000 Basic Word List のレベル 5 までの語彙であることがわかる。ただし、この数字の解釈には語数のカウント方法に影響を受けることに注意が必要である。語彙のカウント方法は大きく、総語数、異なり語数、レマなどがある。総語数は当該コーパスに登場した単語すべての数を示し、異なり語数は綴り字が異なる単語を 1 つずつカウントしたものを指す。例えば、長文の場合、冠詞や前置詞などの機能語が繰り返し登場するため、“I play tennis. I want to be a tennis player in the future.” のような文章では、総語数 (token) は 13 語である。I は 2 回、tennis も 2 回登場しているため、異語数 (type) は 11 語である。表 1・2 に示される割合と累積率は総語数を用いたものであり、コーパスサイズが大きくなればなるほど高頻度語が増えるという Zipf の法則に従うと、解釈に必要な数字である。

表 1 中学校および高等学校検定教科書コーパスに関する情報

コーパス	レベル	異語数		総語数		割合		累積率
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
	固有名詞	407.75	129.52	3167.75	958.68	15.74	1.45	15.74
	L1	641.75	23.13	14951.75	5650.40	72.76	2.35	88.49
	L2	152.50	19.71	784.00	411.78	3.69	0.61	92.18
	L3	66.50	8.50	379.50	67.43	2.05	0.83	94.23
	L4	57.50	9.61	245.75	96.02	1.20	0.25	95.43
中 1	L5	16.00	3.16	61.75	32.36	0.34	0.23	95.76
	L6	11.25	2.87	34.50	17.45	0.18	0.09	95.94
	L7	7.75	1.71	20.75	14.84	0.10	0.05	96.03
	L8	7.75	1.89	49.25	31.45	0.24	0.11	96.27
	NA	108.25	25.84	769.25	323.98	3.74	0.49	100
	ALL	1477.00	157.50	20464.25	7404.84	-	-	-

<sup>1</sup> Mizumoto, A. (2021). New Word Level Checker [Web application]. <https://nwlc.pythonanywhere.com/>

	固有名詞	381.50	102.76	2707.00	1063.67	13.05	1.31	13.05
	L1	745.50	65.08	16045.50	6847.80	76.07	1.55	89.12
	L2	224.75	44.98	1003.75	532.66	4.63	0.40	93.75
	L3	83.25	15.90	350.00	74.96	1.79	0.51	95.54
	L4	59.75	18.48	249.25	143.13	1.13	0.21	96.67
中 2	L5	20.50	6.76	88.75	75.39	0.38	0.15	97.05
	L6	13.00	5.16	40.50	37.25	0.17	0.08	97.22
	L7	13.50	6.14	35.75	17.35	0.17	0.04	97.39
	L8	6.75	3.20	36.50	33.97	0.16	0.09	97.55
	NA	94.25	34.85	543.25	341.49	2.46	0.55	100
	ALL	1642.75	217.22	21100.25	9116.70	-	-	-
	固有名詞	374.25	54.67	2574.75	1009.04	11.49	1.74	11.49
	L1	799.25	55.47	17666.00	8041.40	77.15	1.50	88.64
	L2	315.25	43.22	1307.75	818.45	5.45	0.78	94.09
	L3	115.75	21.69	439.00	152.51	2.00	0.47	96.09
	L4	64.50	11.68	222.25	99.22	0.97	0.04	97.06
中 3	L5	28.75	6.80	87.25	58.52	0.36	0.10	97.42
	L6	17.50	1.29	46.50	23.78	0.21	0.07	97.63
	L7	10.25	3.50	23.00	11.40	0.10	0.01	97.73
	L8	10.50	2.38	37.50	13.77	0.17	0.06	97.9
	NA	96.25	20.45	475.50	236.52	2.09	0.56	100
	ALL	1832.25	171.04	22879.50	10367.23	-	-	-
	固有名詞	211.50	68.59	765.50	94.05	8.42	4.16	8.42
	L1	664.50	205.77	8300.50	5236.13	76.90	3.05	85.32
	L2	310.00	173.95	821.00	656.20	7.13	1.88	92.45
	L3	133.50	72.83	295.00	195.16	2.70	0.21	95.15
	L4	60.00	41.01	112.50	85.56	1.00	0.21	96.15
高 1	L5	37.50	17.68	90.00	35.36	0.91	0.21	97.06
	L6	18.00	8.49	37.00	21.21	0.35	0.01	97.41
	L7	15.50	3.54	41.00	4.24	0.46	0.23	97.87
	L8	12.00	12.73	28.50	31.82	0.22	0.17	98.09
	NA	78.50	3.54	176.50	26.16	1.93	0.91	100
	ALL	1541.00	601.04	10667.50	6385.88	-	-	-

	固有名詞	360.50	137.89	2340.50	1306.03	10.91	4.18	10.91
	L1	794.00	148.49	19290.00	16667.92	74.54	1.58	85.45
	L2	462.50	249.61	1905.50	1830.70	6.89	1.28	92.34
	L3	227.00	182.43	698.00	702.86	2.44	0.66	94.78
	L4	114.50	86.97	338.00	333.75	1.20	0.28	95.98
高 2	L5	70.00	62.23	230.50	252.44	0.76	0.35	96.74
	L6	28.50	19.09	83.00	63.64	0.34	0.04	97.08
	L7	31.50	26.16	79.50	68.59	0.31	0.01	97.39
	L8	23.50	21.92	73.50	62.93	0.29	0.01	97.68
	NA	147.50	99.70	610.50	531.04	2.35	0.07	100
	ALL	2259.50	1034.50	25649.00	21819.90	-	-	-
	固有名詞	238.00	52.33	1167.50	221.32	7.34	3.47	7.34
	L1	774.00	96.17	13959.00	9084.91	74.04	1.51	81.38
	L2	439.00	209.30	1746.50	1578.97	8.33	3.15	89.71
	L3	240.00	156.98	682.50	679.53	3.13	1.65	92.84
	L4	112.00	56.57	266.50	164.76	1.43	0.03	94.27
高 3	L5	68.00	32.53	170.00	111.72	0.90	0.03	95.17
	L6	38.50	16.26	120.00	60.81	0.68	0.11	95.85
	L7	23.00	8.49	57.00	35.36	0.31	0.01	96.16
	L8	21.00	2.83	51.00	5.66	0.35	0.25	96.51
	NA	119.00	36.77	513.00	42.43	3.52	2.46	100
	ALL	2072.50	668.22	18733.00	11889.29	-	-	-

Note. NA は Not Available を指しており、New JACET 8,000 Basic Word List に含まれていない語彙のことを指す。New JACET 8,000 Basic Word List の対応レベルの算出は New Word Level Checker を用いて行われた。

教科書コーパスに収録されている教科書は、令和 5 年 3 月現在に採用されている中学 1 年から 3 年および高校 1 年は新学習指導要領に基づき、高校 2 年および 3 年は旧学習指導要領に基づく。なお、高校の検定教科書は英語コミュニケーションを対象としている。これらは全国の中学・高等学校における採択率の高いものを収集した。具体的には、中学校では Sunshine、NEW HORIZON、NEW CROWN、Here We Go! の 4 種、高校では CROWN、ELEMENT の 2 種を対象とした。コーパス化の範囲に関しては、リーディングパッセージで使われているものだけでなく、リスニングスクリプトや巻末資料のリスト、言語活動の Word Box の注釈などもコーパス化した。なお、指導書とセットの電子データ (e.g., Sunshine の「小学校で習った語彙」や CROWN の新出語彙の英英定義・例文) は除外し、生徒のアクセスできる紙版での教科書に収録されているもののみを対象としている。また、中高教科書コーパスの数値に関しては、複数の教科書を用いて学習する生徒はほとんどいないことを踏まえ、各教科書会社の数値の平均と標準偏差を掲載している。

TOEIC L&R は 2016 年に、TOEIC Bridge L&R は 2019 年にそれぞれアップデートされている。その内容に準拠した問題集は ETS が両テストの制作と同じプロセスで書き下ろしているため、実際のテストをある程度反映していると考えられる。本調査ではこの問題集をコーパス化し、TOEIC L&R および TOEIC Bridge L&R で用いられている語彙について検討する。なお、教科書コーパスと同様に、リスニングスクリプトや設問指示なども収集した。共通テストは令和 3 年度から令和 5 年度に実施されたものを収集し、リスニングスクリプトや設問選択肢を含めてコーパス化した。なお令和 3 年度と 4 年度は追試問題が含まれているが、令和 5 年度は含まれていないことに留意されたい。

表2 TOEIC L&amp;R、TOEIC Bridge L&amp;R、共通テストコーパスに関する情報

コーパス	レベル	異語数	総語数	割合	累積率
TOEIC L&R	固有名詞	2774	4760	9.70	9.70
	L1	959	22328	45.48	55.18
	L2	858	9713	19.79	74.97
	L3	754	5277	10.75	85.72
	L4	558	2362	4.81	90.53
	L5	448	1405	2.86	93.39
	L6	271	827	1.68	95.07
	L7	184	373	0.76	95.83
	L8	129	273	0.56	96.39
	NA	983	1772	3.61	100
	ALL	7918	49090	100	-
TOEIC Bridge L&R	固有名詞	768	4294	11.44	11.44
	L1	802	27678	73.74	85.18
	L2	410	2590	6.90	92.08
	L3	219	937	2.50	94.58
	L4	96	205	0.55	95.44
	L5	49	120	0.32	95.59
	L6	24	54	0.14	95.69
	L7	19	41	0.11	95.77
	L8	19	30	0.08	95.77
	NA	126	1586	4.23	100
	ALL	2532	37535	100	-
共通テスト	固有名詞	956	3332	6.94	6.94
	L1	951	35855	74.73	81.67
	L2	779	4530	9.44	91.11
	L3	550	1730	3.61	94.72
	L4	275	712	1.48	96.2
	L5	139	271	0.56	96.77
	L6	77	198	0.41	97.18
	L7	41	87	0.18	97.36
	L8	42	107	0.22	97.58
	NA	369	1160	2.42	100
	ALL	4179	47982	100	-

Note. 標準化テストの情報はサブコーパスを作成していないため平均値と標準偏差を記載していない。



### 3.2 語彙の多様性の観点からのコーパス情報の解釈

3.1 節で示した表 1 および表 2 は語彙カウントの方法として、総語数と異語数を採用し、New JACET 8,000 Basic Word List のレベル別でどの程度の語彙が用いられているかを示した。総語数換算は語彙の重複を許すという数え方であり、実際の英文理解に近いカバー率を検討することができる。例えば、総語数が 200 words の英文を読解する際、レベル 5 までの語彙の累積カバー率が 95% の場合、残りの約 10 語がレベル 5 を超える語彙であることが分かる。そのため、教科書の 1 レッソンの英文の客観的な難易度を推定する際には有用な指標であると言える。しかし、今回のようにコーパスの総語数が膨大な場合、英文では高頻度の機能語や基本動詞が繰り返し使われ、低頻度の名詞などについては登場回数が少ない。この Zipf の法則のため、高頻度語彙の存在が過剰に高く評価されてしまうことから、レベル 6 を超える難易度の高い語彙の割合は非常に小さく推定されてしまう。そこで、多様な語彙が使われているかどうか（語彙の多様性）を判断の指標とすることが必要である。語彙の多様性はテキスト中の異なる語彙の数を表し、この値が高いほど語彙的に複雑なテキストであると言える。表 3 は表 1 および表 2 に示される New JACET 8000 Basic Word List に含まれている語彙 (i.e., L1~L8) に焦点を当て、語彙の多様性の指標である Type Token Ratio (TTR : 異語数を総語数で割った値)、Guiraud index (異語数を総語数の平方根で割った値)、HD-D (Herdan's C : 異語数を総語数で割り、その商を自然対数にかけた値) を計算したものである。テキストの語彙数が多くなるにつれ TTR の分母は大きくなることから、総語数の影響を受けやすい。この影響を取り除くため、Guiraud index は平方根、HD-D は自然対数で正規化している。これらの値は高いほど、多様な語彙が使われていることを意味している。

表 3 語彙の多様性

	異語数 (Types)	総語数 (Tokens)	Type Token Ratio	Guiraud Index	HD-D
中 1	961.00	16527.25	5.81	7.48	0.71
中 2	1167.00	17850.00	6.54	8.73	0.72
中 3	1361.75	19829.25	6.87	9.67	0.73
高 1	1251.00	9725.50	12.86	12.69	0.78
高 2	1751.50	22698.00	7.72	11.63	0.74
高 3	1715.50	17052.50	10.06	13.14	0.76
TOEIC L&R	4161	49090	8.50	18.78	0.77
TOEIC Bridge L&R	1638	37535	4.4	8.45	0.70
共通テスト	2854	47982	5.9	13.03	0.74

Note. New JACET 8000 Basic Word List に含まれていない NA の語彙や固有名詞を除いて計算している。中高の教科書は平均値を採用している。

$$Type\ Token\ Ratio = \frac{Type}{Token} \times 100, Guiraud\ index = \frac{Type}{\sqrt{Token}}, HD-D = \frac{\ln Type}{\ln Token}$$

概して教科書で用いられている語彙は学年が上がるにつれて高くなる傾向が示された。ただ、高校 1 年から 2 年にかけては減少が見られたことから、学習指導要領の改訂に伴い、指導する語彙が多様化していることが窺えた。各標準化テストの結果は、TTR は数値が低いが、Guiraud index および HD-D は高い値が示された。この理由は Token が大きく、母数が大きくなるためである。語数に比較的頑健な HD-D の結果を参照すると、各標準化テストの語彙の多様性は教科書のそれとほとんど変わらないことが分かる。これらの結果と、3.1 節の結果を合わせて考察すると、TOEIC L&R、TOEIC Bridge L&R、共通テストにおいても New JACET 8,000 Basic Word List における客観的指標を用いて難易度を推定することが可能であると考えられる。つまり、教科書および各標準化テストで扱われている語彙の約 9 割はレベル 5 までの語彙で構成されていると言える。

### 3.3 レマ換算のカバー率

カバー率を考える上では、語彙の重複や異語数の数え方ではなく、レマ換算で計算することが妥当であると思われる。レマ換算とは辞書の見出しのように、屈折を含めた数え方である。例えば、異語数として play と plays は 2 語であるが、レマでは 1 語とみなされる。表 4・表 5 および図 1 は固有名詞を除いた上で、TOEIC L&R、TOEIC Bridge L&R、共通テストの語彙が教科書の語彙でどのくらいカバーされるかの結果を示す (ivy [ツタ] や peony [芍薬] のように普通名詞として用いられるものは含めている)。表 4 は中学校の学年別および累計における各標準化テストの使用語彙のカバー率を表す。表 5 は高校 2 社の検定教科書の平均に基づくカバー率を示している。例えば、コーパス化した 4 社の検定教科書の平均に基づく、中学校全体で TOEIC L&R の 33.90% (Min = 31.3%、Max = 37.0%)、TOEIC Bridge L&R の語彙の 64.02% (Min = 60.2%、Max = 68.6%)、共通テストの 49.21% (Min = 45.6%、Max = 53.4%) がカバーされていることが示された。この数値は、中学生・高校生が複数の教科書を用いて英語を学習する状況が現実的ではないことを踏まえ、1 社のみを対象とした平均値を用いたものである。なお、旧学習指導要領の教科書を用いた卯城・濱田 (2017) では中学校の教科書 (1 社のみ) の計算) では TOEIC L&R の使用語彙の 30% ~ 36% を教科書でカバーしていることが示されている。本調査では新たに TOEIC Bridge L&R と共通テストのカバー率を加えた。

高校卒業時点までのカバー率を付録 6 に示す。ただし中学校では 4 社、高校では 2 社のみをサンプルとしていることから、サンプル数が少なく標準偏差が大きい点に留意されたい。ここでの表記は、例えば中学校で A 社、高校で E 社の教科書を使って学習した生徒が各種標準化テストで使用されている語彙をどのくらい理解できるかを検討している。つまり、B 社に特有な語彙を当該生徒は学習していないという前提に基づいている。この高校卒業までの「全体」の語彙カバー率の平均値として、TOEIC L&R は 54.4%、TOEIC Bridge L&R は 79.1%、共通テストは 72.6% であった。これらの結果から、大学入試を希望する高校生は共通テストの練習として TOEIC Bridge L&R を受験することも視野に入れることができるだろう。また、新学習指導要領に基づいて教科書が改訂されることにより、今後カバー率が高くなる可能性があるかと推測される。

表 4 中学校検定教科書における各標準化テストのカバー率

テスト		中学 1 年	中学 2 年	中学 3 年	中学 1 年-2 年	中学 1 年-3 年
TOEIC L&R	<i>M</i>	17.8	21.9	25.0	26.7	33.9
	<i>SD</i>	1.07	2.37	2.43	2.18	2.82
TOEIC Bridge L&R	<i>M</i>	39.0	46.5	51.2	54.4	64.0
	<i>SD</i>	1.52	4.76	3.80	3.70	3.98
共通テスト	<i>M</i>	26.5	32.6	37.6	39.1	49.2
	<i>SD</i>	1.31	3.53	3.42	3.18	3.90

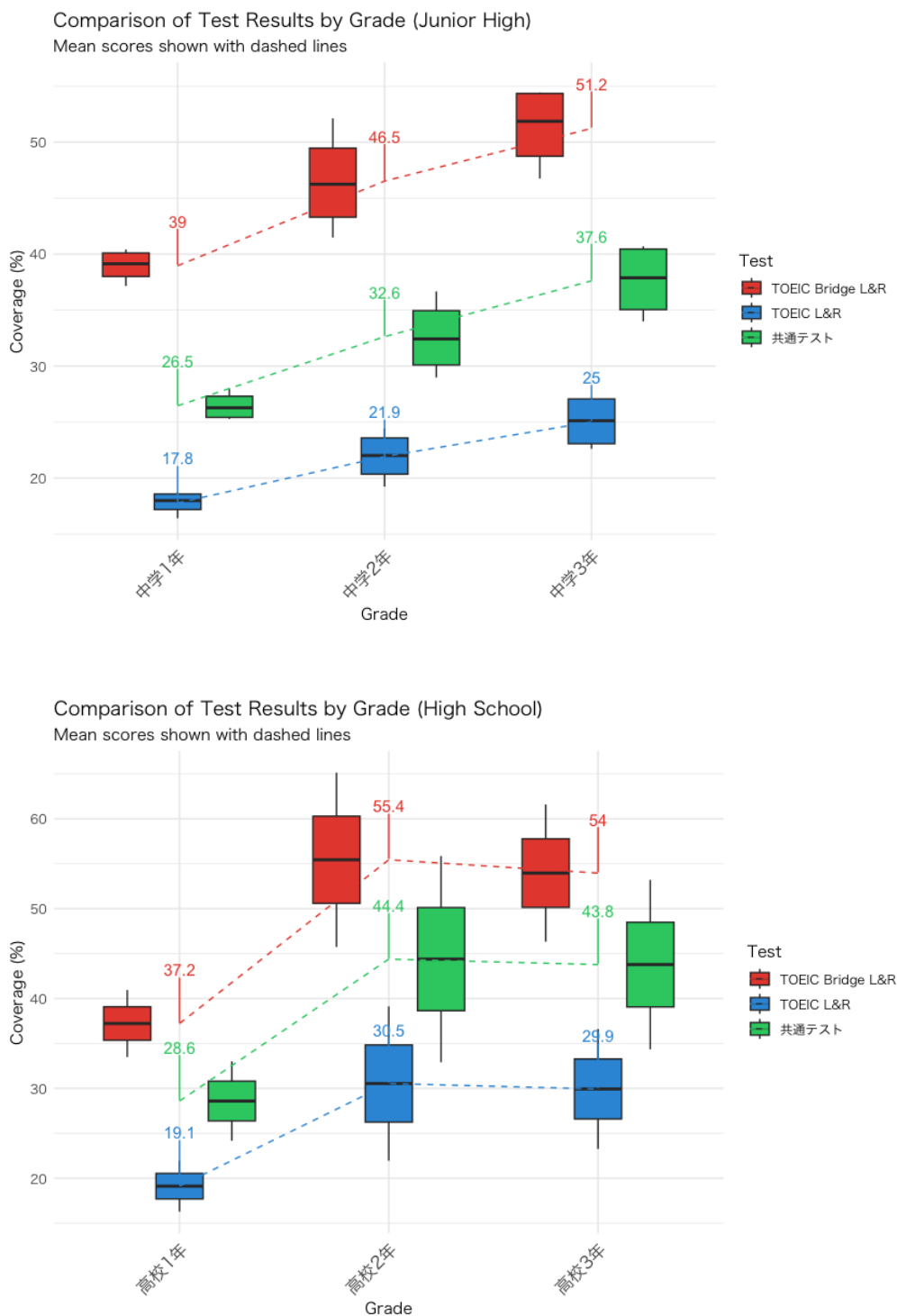
Note. *N* = 4。

表 5 高校検定教科書における各標準化テストのカバー率

テスト		高校 1 年	高校 2 年	高校 3 年	高校 1 年-2 年	高校 1 年-3 年
TOEIC L&R	<i>M</i>	19.1	30.5	29.9	35.7	42.2
	<i>SD</i>	4.01	12.1	9.44	12.2	13.0
TOEIC Bridge L&R	<i>M</i>	37.2	55.4	54.0	61.6	68.5
	<i>SD</i>	5.28	13.7	10.8	12.7	11.9
共通テスト	<i>M</i>	28.6	44.4	43.8	50.7	58.5
	<i>SD</i>	6.25	16.2	13.3	15.6	16.3

Note. *N* = 2。

図1 各学年および累積でのカバー率



教科書は使用されているトピックが多岐に渡り、各標準化テストの対策のためだけに作られているわけではないことから、教科書間で用いられている語彙カバー率の分散が大きいことに留意されたい。なお、中学校・高校で用いられている語彙という大きな枠組みで教科書コース全体から各標準化テストのカバー率を計算した結果は付録4、TOEIC L&R、TOEIC Bridge L&R、共通テストを用いて教科書のカバー率を計算した結果は付録5を参照のこと。なお中高教科書全体で用いられている語彙（レマ換算）を用いると、TOEIC L&Rは60%前半から後半、TOEIC Bridge L&Rは80%以上のカバー率が得られ、共通テストと似た傾向が見られた（付録4参照）。約2%程度であるが、新学習指導要領に基づく教科書でのカバー率は旧学習指導要領に基づく教科書でのカバー率よりも高かったことから、高校2年および3年の教科書が改訂されることにより、今後カバー率はより高くなる可能性がある。また、小学校でふれる600～700語や中高における教科書以外の教室英語も含めるとさらにカバー率は上昇することが予想される。

TOEIC L&R、Bridge L&R、および共通テストから見た教科書の語彙についても、対比のためカバー率を分析した（付録 5）。このカバー率は各標準化テストで使われている語彙が教科書で学習された語彙であることのクロスバリデーションとして機能する。その結果、付録 4 に示す教科書から見たカバー率とほとんど同じであることが示された。この結果は、TOEIC L&R および TOEIC Bridge L&R を受験することにより英語を学びなおす社会人や大学院進学を志す学部生などは、その学習によって、新課程の中高英語教科書のかなりの語彙をカバーすることになり、生涯学習としての意義があると言える。ただし、この数字は各標準化テストで使われている語彙を全て理解した上で、教科書で使用されている語彙をどのくらい理解できるかの割合を意味している点に留意が必要である。さらに、TOEIC L&R や TOEIC Bridge L&R の問題集には出現しなかった語彙があったという結果は、TOEIC L&R を繰り返し受験した大人であっても、中高教科書に掲載されているコミュニケーションの目的・場面・状況に応じた語彙の適切な使用を学ぶ必要性を示している。なお、付録 4 の結果と同様に、旧学習指導要領に基づく高校 2 年生および 3 年生の教科書は新学習指導要領に基づく教科書のカバー率よりも低いことが示された。今後、新学習指導要領に基づいて教科書が改定されることで、教科書で扱われる語彙で各標準化テストの語彙をカバーする割合が高くなる可能性がある。

なお、共通テストと TOEIC L&R および TOEIC Bridge L&R の関係は表 6 に示す通りである。表 6 に示されるカバー率は、共通テストで使用されている語彙を全て理解した上で、TOEIC L&R および TOEIC Bridge L&R で使用されている語彙を理解できるのはどの程度かの解釈につながる。例えば、共通テストのリスニングセクションで用いられている語彙を全て理解していれば、TOEIC L&R のリスニングセクションで使用されている語彙の 45.54%を、TOEIC L&R のリーディングセクションで使用されている語彙の 33.80%を、それぞれ理解できることを意味している。リーディングとリスニングを分けずに、計算したところ、共通テストが TOEIC L&R および TOEIC Bridge L&R の使用語彙をカバーする割合は、TOEIC Bridge L&R で 78.93%、TOEIC L&R で 50.40%であった。TOEIC L&R および TOEIC Bridge L&R の語彙から見た共通テストの語彙カバー率については、TOEIC Bridge L&R で 79.11%、TOEIC L&R で 43.68%であった。リスニング・リーディングの別に関する結果は付録 7 を参照のこと。

表 6 共通テストで使用されている語彙が TOEIC L&R および TOEIC Bridge L&R で使用されている語彙をカバーする割合

テスト	サブコーパス	共通テスト	
		Listening	Reading
TOEIC L&R	Listening	45.54	64.50
	Reading	33.80	53.32
TOEIC Bridge L&R	Listening	60.08	77.85
	Reading	54.32	67.19

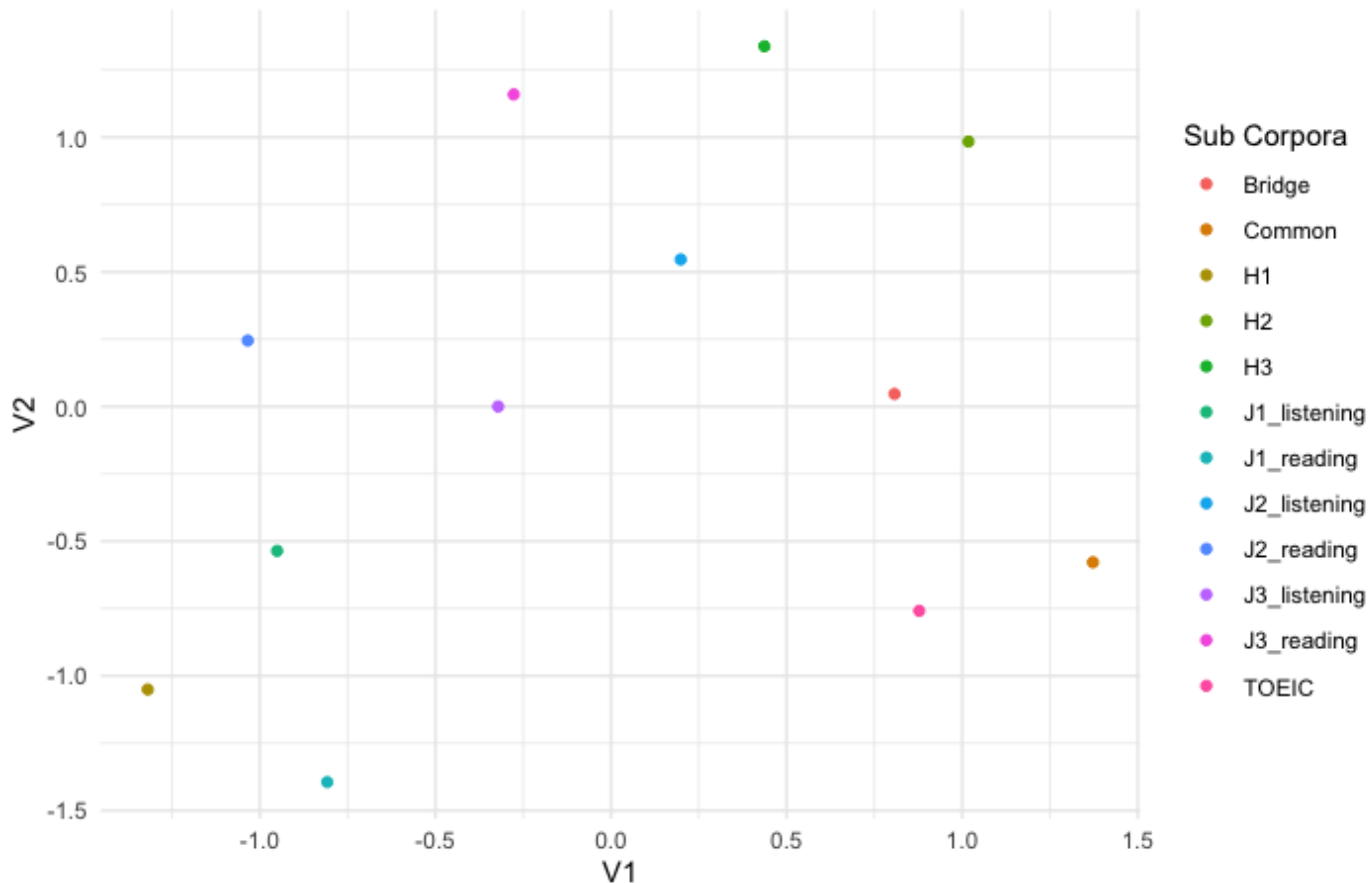
カバー率に関する結果から（表 4、表 5、表 6、付録 4、付録 5、付録 6、付録 7）、教科書で学んだ語彙の観点で相対的に 3 種の標準化テストの難易度を順位づけるのであれば、最も語彙的に易しいテストは TOEIC Bridge L&R、次いで共通テスト、最も難易度が高いのは TOEIC L&R であった。ただし、学習指導要領の改訂に伴い、教科書に収録される語彙の多様性が変化することでこの相対的な難易度は変化することに留意されたい。

### 3.4 中高英語教科書、TOEIC L&R、TOEIC Bridge L&R、共通テストで使用されている語彙の概観

収録語数が多い各種コーパスの傾向を掴むため、本調査では UMAP を用いた。UMAP は次元削減の教師なし学習の機械学習の一種であり、主成分分析のように 2 次元の特徴量空間に観測値を配列することができる。主成分分析では削減する特徴量（変数）が量的である必要がある。一方、機械学習の UMAP ではカテゴリカル・連続の膨大な量の特徴量をモデルに組み込むことが可能である。特徴量として、AntConc で抽出した 4 つのコーパス（TOEIC L&R、TOEIC Bridge L&R、共通テスト、教科書本文）の単語語彙リストおよび、AntGram で抽出した 4 つのコーパスの Trigram を用いて、UMAP を実施した。サブコーパスとして教科書の学年を分けたもの（図 2）とまとめたもの（図 3）の 2 つのモデルを作成した。

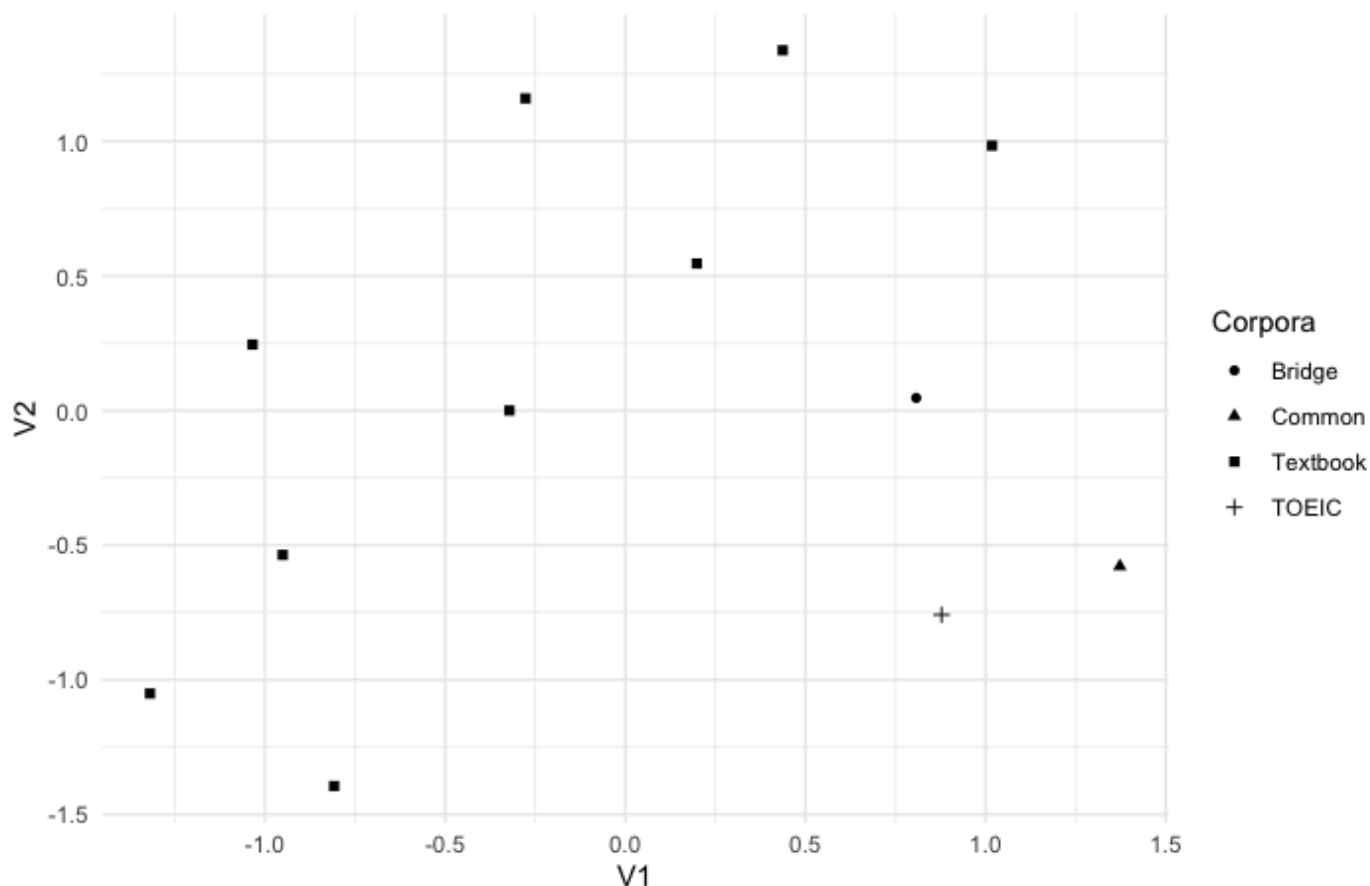
UMAPの2次元ベクトルは名前をつけずに、プロットされている項目の距離を元に解釈することが一般的である。今回は特徴量として、単語および Ngram の出現頻度のみを使っており、文法的難易度や話し言葉のモードなどについては検討していない。2つのプロットからは以下の2点が言える。1点目は TOEIC L&R と共通テストは比較的近くに配列されていることから、共通テストで使われている語彙は TOEIC L&R の語彙と似ている傾向がある点である。2点目は TOEIC Bridge L&R は教科書の近くに配列されていることから、これらのコーパスの語彙が似ている傾向がある点である。2点目は教科書で使われている語彙のカバー率が最も高かったのが TOEIC Bridge L&R であることに裏付けられている。ただし、これらの結果を解釈するにあたり、似ているのは単語レベルなのか、Multi Word Unit のレベルなのかは UMAP では明らかにすることができない。そのため、クロスバリデーションとして、ランダムフォレストとコレスポネンス分析を実施する。

図2 教科書の学年を特徴量空間にプロットした UMAP の出力結果



Note. Bridge は TOEIC Bridge L&R、TOEIC は TOEIC L&R、Common は共通テスト、H は High School、J は Junior High School、数字は学年をそれぞれ表している。

図3 各標準化テストを特徴量空間にプロットした UMAP の出力結果



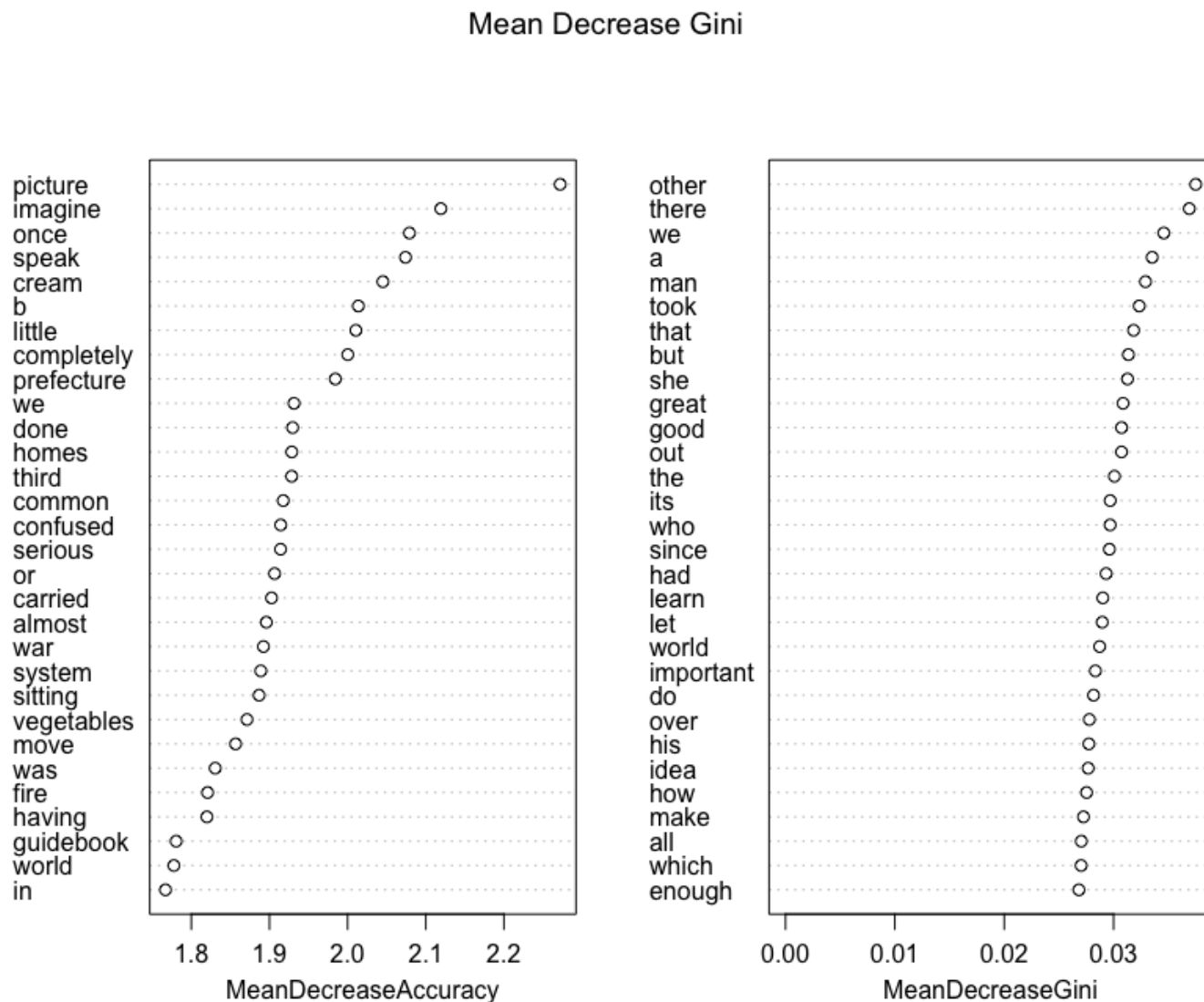
### 3.5 単語・Ngram でコーパスの分類

#### 3.5.1 ランダムフォレストによる単語の出現頻度での分類

4種類のコーパス（TOEIC L&R コーパス、TOEIC Bridge L&R コーパス、共通テストコーパス、教科書コーパス）で用いられている単語の語彙リストから、どの語彙でそれぞれのコーパスを分類可能かについて、ランダムフォレストにより検討した。ランダムフォレストは教師ありの機械学習であり、分類決定木と同様に、目的変数の分類・予測を行うことができる。分類決定木ではデータ数が少ない場合には過学習が起きてしまう。本研究はサンプル数が少ないため、過学習が起きやすい決定木よりもランダムフォレストを採用した。ランダムフォレストの目的はコーパスの分類に寄与する語彙を選定することにある。具体的にはジニ係数減少のグラフを確認することで検討する。

2000本の決定木、16,042の特徴量を用いたランダムフォレストの結果は以下の通りである。ランダムフォレストの目的変数は4種類のコーパスであり、特徴量は単語の出現頻度となる。ジニ係数減少のグラフによると、分類に寄与する重要度は上位から順に、other、there、we、manなどの高頻度語彙であり、難易度の高い低頻度語彙は含まれていない。ジニ係数の減少具合は非常になだらかとなっていることから、単語レベルでのコーパスの分類は難しいように思われる。単語レベルでコーパスの分類が難しく、低頻度語彙が重要度として抽出されなかったという結果は、TOEIC L&R および TOEIC Bridge L&R で用いられている語彙の傾向は教科書で用いられている語彙と似ていることを示している。

図4 ジニ係数減少のグラフ



Note. ジニ係数の減少のグラフは右側である。上から順に目的変数の分類に寄与する特徴量であると判断できる。

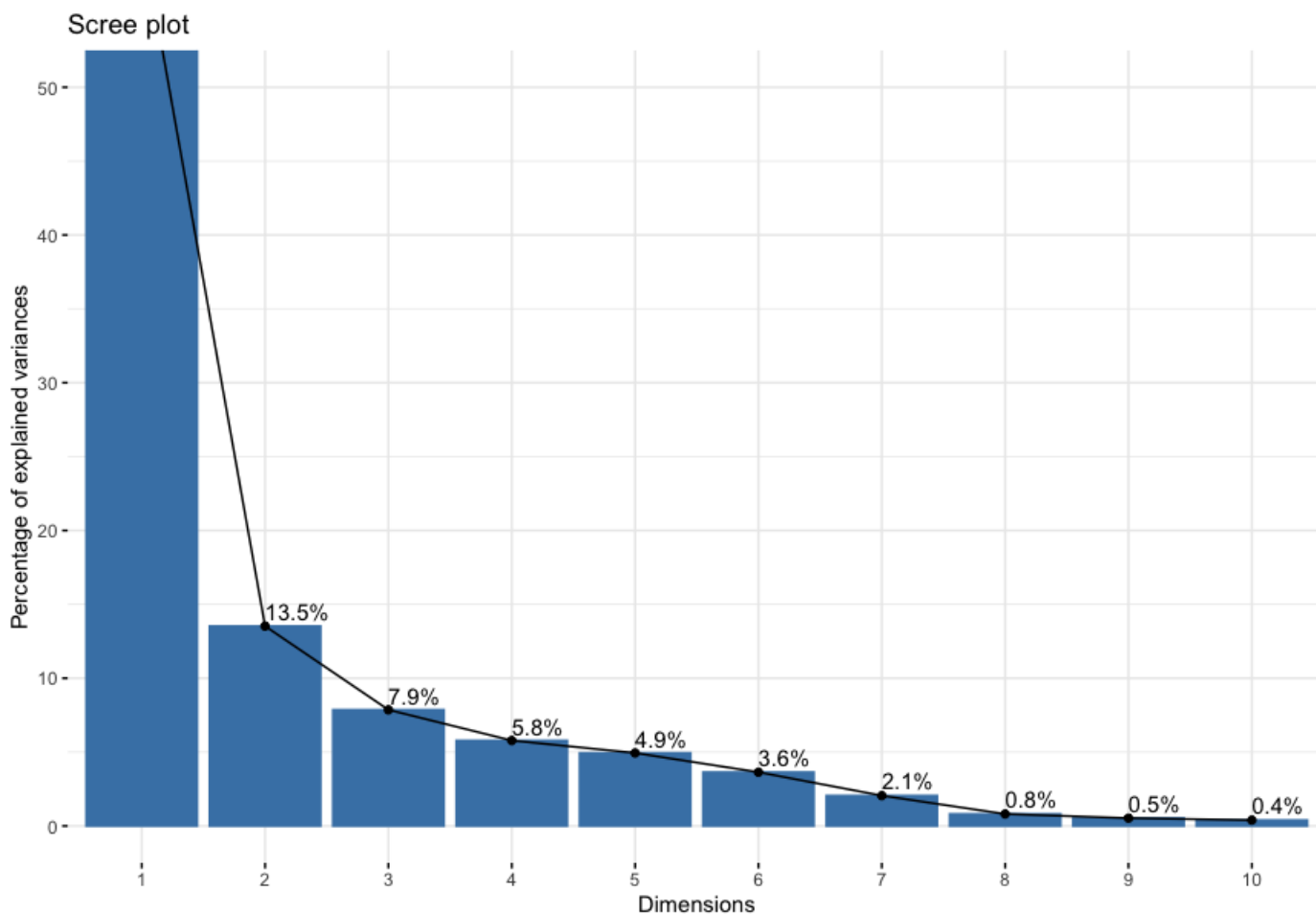
なお、今回はサンプル数が非常に少なく、過学習に陥ってしまった可能性が高い。今後はサブコーパスを複数作成した上で、語彙の難易度や英文自体の読みやすさなどを含めて包括的に検討する必要がある。また、それぞれのコーパスに特徴的な語彙については、より質的な検討が必要となる。

### 3.5.2 コレスポネンス分析による Ngram の出現頻度での分類

UMAP では大量の特徴量を組み込んだモデリングを行ったが、個々の Ngram との関係性については検討していない。そこでコレスポネンス分析を用いて、出現頻度が上位 50 項目に属する Ngram (付録 3 参照) を同一平面上にプロットすることで、4 つのコーパスの傾向を調査する。コレスポネンス分析とは、多数の質的変数同士の連関について、多次元空間上でのカテゴリの対応関係を 2 次元空間に布置する手法である。UMAP と同じ次元圧縮であるが、コレスポネンス分析はカテゴリを平面で視覚的に捉えることができる点で優れている。

Ngram を列項目、4 つのコーパスを行項目としたコレスポネンス分析の結果、10 の次元が抽出された。

図 5 固有値に基づく次元のスクリープロット

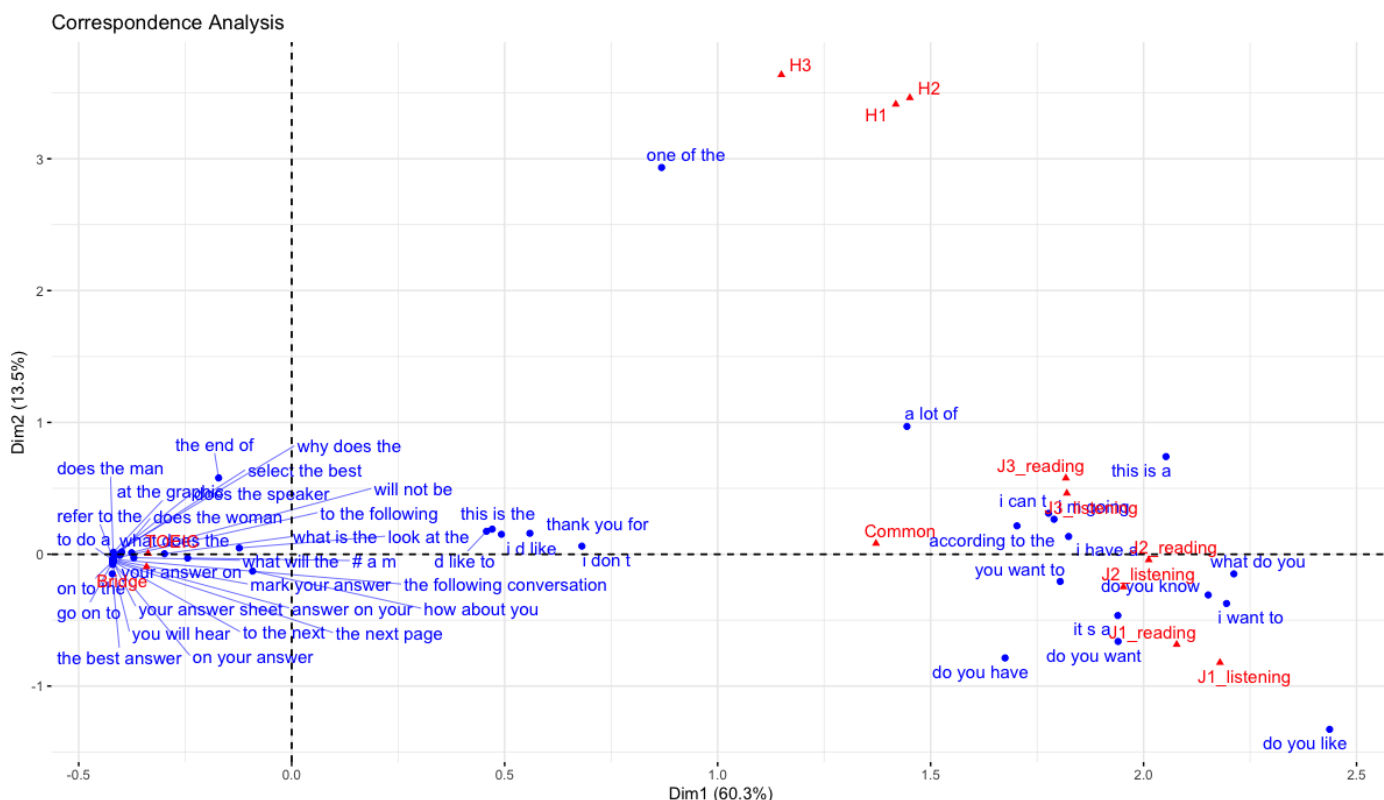


Note. X 軸には抽出された次元の数、Y 軸には次元ごとの固有値をプロットしている。この固有値はイナーシャの寄与率と呼ばれる。

次元 1 と次元 2 のイナーシャの寄与率はそれぞれ 60.3%、13.5%であった。プロットを確認すると、Ngram の観点からは、高校教科書 (H1~3) は他のコーパスとは独立してまとめられており、中学校と共通テストの周囲には want を中心とした文法的コロケーションや助動詞 can, do you の Yes/No 疑問文などがまとまっていた。TOEIC L&R と TOEIC Bridge L&R の問題集は多くの Ngram が集まっていることから、連語の観点からよりオーセンティックなマテリアルを用いていると考えられる。



図 6 コレスポネンス分析の出力結果



ランダムフォレストおよびコレスポネンス分析の結果をまとめると、単語レベルでは 4 種類のコーパス (中高検定教科書・TOEIC L&R・TOEIC Bridge L&R・共通テスト) を分類することが難しいことが示された。コレスポネンス分析では中高の検定教科書を学年のサブコーパスに分け、Ngram の観点から (a) 中学校 1 年から 3 年の教科書と共通テスト、(b) 高校 1 年から 3 年の教科書、(c) TOEIC L&R と TOEIC Bridge L&R はそれぞれ類似していることを示した。

### 3.6 中高英語教科書および共通テストから見た TOEIC L&R に特有な語彙

3.5 節ではデータ全体を用いて、それぞれのコーパスの傾向を掴んだ。ランダムフォレストの結果によると、TOEIC L&R や TOEIC Bridge L&R、共通テストで用いられている低頻度の語彙は教科書の語彙との分類に寄与していない。そのため、それぞれのコーパスで使用されている語彙の特徴をより詳細に検討する必要がある。この手法として、中條・Genung (2005) では、BNC を用いて TOEIC L&R に特有な語彙について以下の 4 つの統計指標 (中條・内山、2004) で、TOEIC L&R に対応するための単語リストを作成している。

- **ダイス係数**：高頻度語が抽出されやすいことから、基本語彙の復習が必要な学習者向け。
- **対数尤度比**：ダイス係数で選択される特徴語よりも一段階上の頻度にある語が抽出されやすいことから、初級レベルの学習者向け。
- **コサイン・補完類似度**：対数尤度比で抽出される単語より少し高レベルの TOEIC L&R 特徴語や英語圏の文化背景とともに教えた日常語など中級レベルの学習者向け。
- **自己相互情報量**：低頻度語が過大評価される傾向にあるもの、意味的な情報量の多い低頻度語を抽出できることから、上級レベルの学習者向け。

本研究では中條・Genung (2005) および卯城・濱田 (2017) と同じ手順に従うが、以下の2点で先行研究と異なる：(a) New JACET 8000 Basic Word List に基づいてレベル別に特徴語を抽出した、(b) 新学習指導要領に基づく教科書に加えて共通テストから単語リストを作成した。表7から表10は、教科書と共通テストを参照コーパスとして、TOEIC L&R および TOEIC Bridge L&R に特有な語彙 (内容語) の上位10語をまとめたものである。

表7 令和4年度新課程教科書に登場する TOEIC L&R で使用頻度の高い語彙

	基本レベル (ダイス係数)	初級レベル (対数尤度比)	中級レベル (類似度)	上級レベル (自己相互情報量)
Level 1				
1	answer	read	example	fill
2	question	part	set	decide
3	go	listen	time	experience
4	have	page	message	lead
5	best	choose	please	save
6	test	example	see	report
7	next	ticket	buy	technology
8	letter	four	thank	system
9	look	short	flower	department
10	hear	talk	train	professional
Level 2				
1	sheet	provide	sample	ship
2	following	screen	painting	offer
3	refer	heat	comment	accord
4	speaker	detail	announcement	financial
5	direction	pack	price	advance
6	conversation	talent	complete	construction
7	schedule	pay	path	policy
8	response	ship	kilometer	additional
9	statement	package	chain	participate
10	text	concern	passenger	recent
Level 3				
1	mark	snack	comedy	client
2	phrase	resident	lunchtime	contrast
3	sentence	cancel	rainy	delivery
4	listener	dancer	reserve	grand
5	section	zoo	exhibit	feedback
6	reserve	instant	pond	paragraph
7	advertise	parade	polish	multiple
8	comfort	glove	parade	reception
9	repair	enjoyable	cafeteria	expense

10	cancel	waiter	discount	annual
Level 4				
1	hike	turkey	cinema	trial
2	mechanic	chess	portrait	inexpensive
3	okay	rider	wildness	inspector
4	registration	assess	printer	downtown
5	café	specially	spice	storage
6	update	guidebook	fisherman	laptop
7	cloudy	comfortably	skate	fitness
8	separately	strip	hint	hardware
9	instruct	holder	sue	medication
10	poetry	pleasant	balloon	outstanding
Over				
Level 5				
1	aisle	builder	rubbish	password
2	opera	tenant	bulb	catalog
3	clothing	sunset	chin	supplier
4	headphone	harbor	backyard	attachment
5	directory	gardener	purely	academy
6	copyright	salon	cord	pharmacy
7	trademark	mar	renewal	logo
8	auditorium	scarf	sin	developer
9	saw	dental	athletics	subscribe
10	logo	prescription	madam	accountant

表 8 令和 4 年度新課程教科書に登場する TOEIC Bridge L&R で使用頻度の高い語彙

	基本レベル (ダイス係数)	初級レベル (対数尤度比)	中級レベル (類似度)	上級レベル (自己相互情報量)
Level 1				
1	close	hand	read	dollar
2	write	wear	question	sound
3	fix	hour	answer	bird
4	have	study	test	parent
5	test	sport	look	neighbor
6	please	vegetable	next	leader
7	read	state	each	add
8	man	window	part	brush
9	ask	thank	hear	farmer
10	help	run	time	tourist

Level 2				
1	provide	passenger	sheet	direction
2	offer	employee	following	kilometer
3	pay	laboratory	direction	allow
4	detail	vehicle	refer	sweater
5	screen	statement	speaker	boot
6	ship	display	conversation	detail
7	concern	suggest	schedule	pound
8	heat	announce	statement	stamp
9	package	patient	response	slow
10	talent	prefer	text	accept
Level 3				
1	reserve	dancer	mark	permit
2	advertise	register	phrase	glove
3	comfort	snack	sentence	indicate
4	mark	receipt	reserve	tire
5	phrase	resident	section	participant
6	sentence	grocery	cancel	decoration
7	listener	slight	discount	spill
8	section	author	snack	seek
9	repair	waiter	advertisement	bush
10	cancel	discount	register	attraction
Level 4				
1	hike	update	café	instruct
2	mechanic	heater	registration	appliance
3	registration	refund	instruct	rear
4	update	manual	update	portrait
5	instruct	hanging	hike	sculpture
6	separately	chef	separately	baker
7	cloudy	chess	cloudy	strip
8	poetry	spice	rear	holder
9	printer	specially	portrait	obey
10	chess	tasty	poetry	hint

Over				
Level 5				
1	aisle	beverage	headphone	headphone
2	opera	tenant	aisle	trademark
3	logo	payroll	directory	bulb
4	auto	homemade	trademark	marker
5	rubbish	gardener	bulb	inspect
6	caller	prescription	opera	whip
7	dental	salon	builder	inning
8	bulb	chin	copyright	cord
9	builder	whichever	logo	pastry
10	tenant	mar	saw	vine

表9 共通テストに登場する TOEIC L&R で使用頻度の高い語彙

	基本レベル (ダイス係数)	初級レベル (対数尤度比)	中級レベル (類似度)	上級レベル (自己相互情報量)
Level 1				
1	people	new	question	mother
2	time	work	school	wild
3	school	most	student	boy
4	get	make	people	castle
5	student	need	choose	farmer
6	like	take	life	born
7	best	now	friend	girl
8	new	next	time	laugh
9	use	let	thing	symbol
10	choose	know	use	uncle
Level 2				
1	speaker	speaker	following	southern
2	following	schedule	yard	tomato
3	type	likely	option	human
4	presentation	online	elephant	horse
5	option	according	based	classmate
6	article	available	ocean	survive
7	schedule	policy	sugar	homework
8	based	equipment	measure	funny
9	complete	price	classmate	silent
10	final	material	survive	scientific

Level 3				
1	carbon	discount	carbon	carbon
2	poster	participant	onion	carrot
3	length	workshop	passage	tape
4	television	recommend	found	creature
5	electronic	arrival	recycle	elephant
6	unit	mark	honey	passage
7	section	unit	slide	shelter
8	author	description	artificial	false
9	slide	indicate	soil	intelligent
10	container	correctly	container	dolphin
Level 4				
1	wallet	appliance	oral	oral
2	documentary	delighted	coral	coral
3	storage	laptop	ghost	fog
4	fog	hike	wallet	sue
5	sue	maple	medal	ghost
6	habitat	accessory	documentary	interval
7	hiking	newsletter	fog	medal
8	draft	participate	sue	ecosystem
9	ghost	upgrade	habitat	slim
10	audio	inexpensive	frustrate	cinema
Over				
Level 5				
1	cabbages	password	gig	gig
2	underwater	warranty	underwater	cabbage
3	saw	academy	patent	underwater
4	owl	seafood	owl	patent
5	flake	receptionist	flake	hokey
6	reef	photography	dormitory	owl
7	tutor	lightweight	gut	flake
8	warranty	digest	reef	gut
9	layout	helmet	ace	reef
10	aquatic	listing	slippery	aquatic

表 10 共通テストに登場する TOEIC Bridge L&R で使用頻度の高い語彙

	基本レベル (ダイス係数)	初級レベル (対数尤度比)	中級レベル (類似度)	上級レベル (自己相互情報量)
Level 1				
1	question	student	student	become
2	time	use	make	bottle
3	school	study	take	collect
4	get	school	people	thing
5	student	however	school	foot
6	like	life	get	mountain
7	new	even	time	pet
8	work	system	question	show
9	choose	international	new	tooth
10	than	get	work	tomato
Level 2				
1	following	sheet	presentation	option
2	final	based	option	elephant
3	item	type	speaker	standard
4	complete	based	type	beans
5	based	presentation	standard	model
6	schedule	final	elephant	drawing
7	article	condition	final	ocean
8	option	result	article	yard
9	presentation	unite	being	carpet
10	type	focus	based	tape
Level 3				
1	length	mark	carbon	carbon
2	poster	poster	length	onion
3	carbon	length	poster	soil
4	slide	television	unit	organic
5	passage	electronic	television	artificial
6	author	grand	onion	summary
7	section	slightly	summary	honey
8	electronic	violin	honey	container
9	unit	valley	container	slide
10	television	recipe	slide	passage

Level 4				
1	wallet	wallet	coral	documentary
2	oral	sue	oral	coral
3	fog	storage	sue	oral
4	shortage	click	habitat	medal
5	documentary	hike	draft	interval
6	audio	downtown	storage	ghost
7	hiking	participation	fog	audio
8	sue	inexpensive	documentary	habitat
9	habitat	holder	wallet	draft
10	draft	appliance	medal	fog
Over				
Level 5				
1	gig	thoughtful	gig	gig
2	saw	mar	cabbage	underwater
3	underwater	curtain	underwater	cabbage
4	cabbage	academy	patent	hockey
5	owl	password	owl	patent
6	hockey	riverside	hockey	flake
7	flake	striped	flake	dormitory
8	concussion	headphone	dormitory	layout
9	gut	beverage	layout	gut
10	layout	attachment	gut	owl

高校を卒業するまでに指導する語数は約 5,000 語であることを踏まえると、5,000 語レベルを超えた語彙は高校を卒業後に自律的に学習する必要がある。5,000 語レベルを超えた低頻度で難易度の高い語彙はどのような傾向があるのかを質的に検討したところ、以下の点が明らかとなった。

教科書および共通テストにはあまり見られない一方で、TOEIC L&R および TOEIC Bridge L&R には頻出の語彙 (New JACET 8,000 Basic Word List でレベル 5 以上の語彙) の傾向は、多くの人が持つ「TOEIC L&R はビジネスパーソン向けのテストである」という印象からは予想しにくいものであった。具体的には、一般的な会話などでも用いられる場所の語彙 (e.g., aisle, directory, riverside, dormitory)、職業の語彙 (e.g., builder, gardener, dental, salon)、商標・著作権の語彙 (e.g., trademark, copyright, patent) などであった。これらの語彙は教科書にも登場しているが、教科書での登場回数は少数に留まっていることが示された。

3.3.節の表 4 および 5 を参照すると、TOEIC L&R および TOEIC Bridge L&R で用いられている語彙に低頻度の難しい語彙が多いというわけではない。むしろ、TOEIC L&R の語彙は教科書を基本としてより低頻度な語彙を広く射程としている標準化テストであると言えるだろう。さらに、抽出された語彙は TOEIC L&R と TOEIC Bridge L&R および共通テストで共通しているものもある。そのため、受験対象を初中級者としている TOEIC Bridge L&R はビジネスパーソンだけでなく、中高生も含めて幅広く受験対象であることが分かる。

表 7 から表 10 の 4 つの表はそれぞれのコーパスに登場した語彙に基づいて抽出された特徴語であるため、最低 1 度は教科書・共通テストに登場している語である。教科書・共通テストに 1 度も登場していないが、TOEIC L&R および TOEIC Bridge



L&R に登場している語彙のリストは付録 1 および付録 2 を参照のこと。具体的な数字は TOEIC L&R に登場し教科書に登場しない語彙は 872 語、TOEIC Bridge L&R に登場し中高の教科書には登場しない語彙は 139 語であった。TOEIC L&R の語彙で特に特徴的なものには、キャリアの語彙 (e.g., candidate, appointment, employer, leadership, negotiate)、電気関係の語彙 (e.g., electronics, cable, component, convert)、典型的なビジネスの語彙 (e.g., promotional, prospect, prospectus, provider) があつた。一方、TOEIC Bridge L&R の語彙で特に特徴的なものには、電子機器の語彙 (e.g., electronics, laptop, wireless)、部屋や場所の語彙 (e.g., auditorium, riverside, roommate, villa) があつた。TOEIC Bridge L&R では accountant や payroll などの働くことに関連する語彙も登場するが、日常的に使われる語彙が多く用いられている傾向が見られた。

#### 4. 結論

本研究は新課程の中高英語教科書、TOEIC L&R、TOEIC Bridge L&R、共通テストの 4 種類をコーパスとし、(a) 中高英語教科書による TOEIC L&R 語彙のカバー率・各種コーパスの語彙が教科書の語彙をカバーする割合を計算、(b) 単語および Ngram の使用頻度からコーパスの分類および (c) 中高英語教科書とは異なる TOEIC L&R 特有の単語の抽出を行った。

令和 4 年度の検定教科書 (高校 2 年・3 年は平成 23 年度施行の学習指導要領に基づく) は、中学校 4 社・高校 2 社の全体を合わせると、TOEIC L&R の語彙の約 71%、TOEIC Bridge L&R の約 91%、共通テストの約 88%を、それぞれカバーしていた (付録 6 に示される平均に基づく結果では、TOEIC L&R は 50.69%、TOEIC Bridge L&R は 79.15%、共通テストは 68.91%)。この高いカバー率の理由は、教科書のアウトプット活動で用いられる tool box や word bank などの注釈、巻末に掲載されている word net や word web などの語彙のネットワークなども含めていることが考えられる。つまり、本研究で用いた教科書は収録されている全ての語彙を計算に含めた結果であるため、教科書本文だけに限定した際のカバー率はより低くなる可能性がある点には留意されたい。ただし、コーパス化した教科書会社間で共通した語彙を使用しているわけではなく、扱われているトピックがバラバラであるにもかかわらず、カバー率は非常に高い値であった。なお、本調査は令和 4 年度時点において出版されている検定教科書を利用していることから、高校 2 年次および 3 年次の新課程教科書をコーパス化していない。今後、全学年の教科書が新課程のものとなった段階ではカバー率は増加する可能性がある。また、教科書で学習した語彙の観点から 3 つのテストの難易度を相対的に比較すると、大学入試を志望する者が受験する共通テストよりも、TOEIC Bridge L&R は易しい傾向が見られた。このことから、特に大学への進学を希望する高校生は積極的に受験することで、共通テストの語彙的難易度の傾向を掴むことができるだろう。語彙指導については、例えば中高生が教科書で触れたことのない未知の語彙が含まれていると推定される TOEIC Bridge L&R の問題集を用いて、未知語に対応するための具体的な方略を指導することも効果的であろう。その際は、TOEIC L&R および TOEIC Bridge L&R に特有な語彙をまとめた付録 1 および 2 を参照されたい。

単語に基づくコーパスの分類は、低頻度で難易度の高い語彙がコーパスの弁別に寄与しているわけではないことが示された。この点は、TOEIC L&R・TOEIC Bridge L&R・共通テストのいずれの標準化テストの受験においても、教科書に登場する基本語彙の習熟が重要であることを意味している。さらに Ngram に基づく傾向は、TOEIC L&R および TOEIC Bridge L&R で Ngram が多く用いられていることから、連語や定型表現を活用したオーセンティックな材料であると言える。Ngram をはじめとした連語や定型表現は (a) 学習者の言語処理速度を高める点、(b) 語彙ネットワークを構築する点、(c) 所属をはじめとするアイデンティティを表す機能がある点などで、近年多くの研究者に注目されている。例えば、according to the のような Ngram はアカデミックな場面でも用いられる定型表現であり、使用者の学術的なアイデンティティを示す際に役に立つ。連語表現が多く収録されている TOEIC L&R および TOEIC Bridge L&R の問題集に取り組むことで、学習すべき語彙に遭遇する頻度を高めることができ、オーセンティックなテキストを読解・聴解する機会を得ることができるだろう。

なお本研究は語彙の観点から TOEIC L&R、TOEIC Bridge L&R、共通テストの傾向を扱ったため、標準テストとしての難易度などについて評価を行うことはできない。つまり、これらのテストが本来測定している聴解および読解能力の関係性については未検討である。また、コーパスの種類が少なかったことから分析が過学習に陥ってしまった可能性があり、サブコーパスを用いてサンプル数を増やした見当も必要となる。今後はこれらの限界点を改善した上で、より詳細な分析が求められる。具体的には、旧学習指導要領下で作成されていた高校 2 年・3 年の教科書コーパスの修正や共通テストコーパスの増加、小学校の教科書コーパス

の分析などが挙げられる。新学習指導要領では高校卒業までに約 5,000 語の語彙を指導することとなるため、約 71%のカバー率であった TOEIC L&R のカバー率も、今後より高くなると予想される。コーパスの語数が増えることにより、それぞれのテストに特徴的な語彙の傾向が変わることも考えられる。さらに、今後はスピーキングおよびライティングで必要となる語彙の傾向についても調査し、どのような語彙を使用する学習者が高得点となるかの予測やタスクごとに用いられているコロケーションが異なるのかの調査などが展望される。

## 引用文献

- Anthony, L. (2022). AntConc (Version 4.2.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Anthony, L. (2022). AntWordProfiler (Version 2.0.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Anthony, L. (2021). AntGram (Version 1.3.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Committee of Revising the JACET Basic Words. (Ed.). (2016). The new JACET list of 8,000 basic words. Kiriara Shoten.
- Hamada, A., Iso, T., Kojima, M., Aizawa, K., Hoshino, Y., Sato, K., Sato, R., Chujo, J., & Yamauchi, Y. (2021). Development of a vocabulary size test for Japanese EFL learners using the New JACET List of 8,000 Basic Words. *JACET Journal*, 65, 23–45. [https://doi.org/10.32234/jacetjournal.65.0\\_23](https://doi.org/10.32234/jacetjournal.65.0_23)
- Kaneko, M. (2017). Vocabulary size targets for the TOEIC test. *JACET journal*, 61, 57–67. [https://doi.org/10.32234/jacetjournal.61.0\\_57](https://doi.org/10.32234/jacetjournal.61.0_57)
- Mizumoto, A. (2021). New Word Level Checker [Web application]. <https://nwlc.pythonanywhere.com/>
- Nation, I. S. P. (2012). *The BNC/COCA word family lists*. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge university press.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics*, 168(1), 33–69. <https://doi.org/10.1075/itl.168.1.02web>
- Zhang, S., & Zhang, X. (2022). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research*, 26(4), 696–725. <https://doi.org/10.1177/1362168820913998>
- 相澤一美・望月正道. (2010). 『英語語彙指導の実践アイデア集-活動例からテスト作成まで』. 大修館書店
- 卯城祐司・濱田彰. (2017). 『新課程の英語教科書コーパスに基づく TOEIC®語彙の特徴と難易度の推定』. 一般財団法人国際ビジネスコミュニケーション協会
- 中條清美・Michael Genung. (2005). 「British National Corpus を活用した TOEIC テストの分析：語彙活用水準の定量化と特徴単語の抽出」『TOEIC Research Report』, 3号, 1–20.
- 中條清美・内山将夫. (2004). 「統計的指標を利用した特徴語抽出に関する研究」『関東甲信越英語教育学会紀要』, 18号, 99–108. [https://doi.org/10.20806/katejo.18.0\\_99](https://doi.org/10.20806/katejo.18.0_99)
- 水本篤. (2006). 「語彙サイズテストは何を測っているのか?—語彙サイズテストの開発における問題点—」『統計数理研究所共同研究レポート』, 190 巻, 71–80. <http://hdl.handle.net/10112/12980>
- 望月正道・磯達夫・高本香織・佐藤繭香. (2011). 「TOEIC リーディング得点を説明する要因の究明」『麗澤レビュー』, 17 巻, 22–41.

付録1 TOEIC L&R に登場し教科書に登場しない語彙

Level 2

additional	appointment	counter	enormous	lane
membership	monthly	previously	rural	sharply

Level 3

absence	additionally	adequate	allergy	analysis
analyst	annually	appoint	aside	assembly
assure	avenue	beneficial	briefly	cable
candidate	collector	competitive	component	comprehensive
consult	convert	coupon	credit	criticism
currently	debt	depart	departure	deposit
dispose	distribute	editor	electronics	employer
enhance	ensure	establishment	evaluate	flexible
foster	frank	install	interaction	invest
lawn	leadership	maintenance	migrate	minimum
minor	motor	negotiate	notable	numerous
orientation	pace	precise	precisely	preference
producer	psychological	publisher	random	reduction
reference	regional	reliable	replacement	reputation
requirement	resemble	retain	retirement	rocky
seminar	significantly	substantial	superior	transition
urgent	via	vital	widespread	wire

Level 4

accommodate	accommodation	actress	administer	administrator
affordable	agenda	alternate	alternatively	appliance
assess	assessment	audio	availability	bargain
boost	brake	bulletin	bypass	cargo
carrier	casual	cater	certificate	certify
chamber	cite	classify	compensate	compensation
compile	composition	considerably	consistently	context
contractor	controversial	convention	corporation	cosmetic
costly	currency	dairy	dare	deadline
diploma	disruption	dominate	doughnut	downstairs
draft	dynamic	economically	economist	edition
effectiveness	elegant	enrollment	entitle	equip
estate	expire	faculty	familiarity	fitness
formally	forum	guidance	halt	harsh
inexpensive	initially	initiative	inspector	instruct

(Level 4)				
insufficient	intensive	involvement	isolate	lag
laptop	lease	literacy	literary	managerial
meantime	mechanic	merchant	merge	migration
moderate	motorcycle	newsletter	outgoing	output
oversee	participation	partnership	penetrate	phase
placement	profile	profitable	prominent	promotion
prospect	protective	provider	publicity	qualification
questionnaire	readily	recruitment	reluctant	reopen
residential	respectively	respondent	retail	revenue
rider	risky	rubber	sector	semester
shorten	signify	simplify	specimen	spelling
stiff	strain	strategic	stressful	supervise
suspend	technician	terrain	textile	timing
transaction	tremendous	unchanged	uncover	unfriendly
unreasonable	update	upgrade	vague	vegetarian
venue	vice	virtually	willingness	wireless

Level 5

abrupt	abruptly	abundance	accent	accordance
acquisition	activate	adequately	adverse	advisory
airway	aisle	alley	allocate	aluminum
amid	anonymous	apology	apparatus	applicable
applicant	appropriately	archaeological	architectural	archive
attorney	auction	authentic	authorize	auto
beverage	bid	binding	breakdown	builder
bureau	calculation	caller	cane	catalog
closure	collective	commence	commuter	compact
comparable	compatible	compliance	complication	compliment
comply	composer	confer	confirmation	congestion
consultation	continuity	contributor	conversion	coordinator
copper	correlation	counsel	counseling	courtyard
coverage	critically	dealer	defective	dental
devise	digit	disappointment	distinction	donor
drawback	duration	editorial	elaborate	eligible
elite	endorse	evidently	exclusively	exotic
expel	expenditure	exposure	extensively	flatten
foam	fond	formulate	forthcoming	fraction
gallon	graphic	identification	implementation	inclusion

(Level5)				
indicator	infrared	infrastructure	intact	intensify
interactive	intermediate	intersection	introductory	inventory
investor	knowledgeable	legendary	lesser	likelihood
likewise	longtime	mandate	mandatory	marble
merger	millennium	minimal	mistakenly	monetary
municipal	narrowly	neatly	nomination	nonetheless
nonprofit	notably	notify	observer	odor
operational	opt	optical	organizational	outdated
outlook	overlap	partial	password	pedestrian
pharmaceutical	pharmacy	pickup	pill	populate
portfolio	practitioner	precision	predecessor	predictable
preliminary	premise	probability	probe	progressive
prospective	protocol	provoke	purely	quota
ram	ratio	reassure	recipient	reconstruction
refine	regulatory	rehabilitation	rehearsal	reinforce
relaxation	reliability	reliance	renewal	renovate
retailer	revision	rim	rotation	rubbish
salon	scent	segment	serial	shareholder
shed	shield	simultaneously	sin	slot
sodium	solely	spacious	structural	submission
substantially	suburban	sufficiently	supervision	supervisor
supportive	surge	symphony	symptom	synthetic
tactic	tenant	therapist	therapy	tile
timber	tolerate	transcript	transformation	turnover
undertake	unveil	upcoming	utility	utilize
vacant	variable	vigorous	violation	ward
warehouse	witty	workforce	workload	workout

Level 6

acclaim	accrue	acoustic	acquaint	align
altitude	approximate	assert	astonish	auditorium
autobiography	automotive	backing	badge	banquet
batch	bin	blur	bookshop	borough
bulk	centrally	certification	chancellor	charcoal
clearance	communicative	confidentiality	contingency	continual
cookbook	copyright	correspond	courtesy	creek
cuisine	customary	defect	demographic	diagram
differentiate	dimension	dismantle	dissatisfied	diversify

(Level 6)				
doorway	drainage	earnings	electronically	elevate
elm	enclosure	enhancement	enthusiast	excerpt
excite	exclusion	experimentation	fabricate	faulty
firsthand	fiscal	fishery	flatter	flaw
fleet	forestry	friction	generalize	handbag
handbook	harden	hardy	healthcare	hesitant
holly	hourly	inaugural	inexperienced	informational
informative	insider	intern	jumbo	justification
landlord	leakage	lightweight	luncheon	massage
meadow	memoir	merchandise	moderately	neutral
offset	ordinance	overdue	paperwork	partition
pat	patiently	patron	pavement	perfume
periodic	persistence	persuasive	pertinent	pest
photocopy	plaque	pleasantly	porch	portal
prescribe	privately	projection	projector	promotional
quarterly	quartet	radius	ranger	receptionist
recreational	redesign	relocation	renovation	renowned
reorganization	restructure	ridge	riverside	roommate
rust	scenic	secondhand	shipment	shoreline
showcase	sizable	slogan	solicit	specialization
specification	spokesperson	straightforward	subscribe	subscriber
subscription	tailor	terrific	thermal	timely
tribute	trouser	turbine	unavailable	undertaking
unload	unpopular	ventilation	verification	villa
voucher	warranty	wholesale	wholly	woodland

Level 7

accountable	accountant	aging	ample	anticipation
apt	armchair	assurance	audit	backup
banner	bedside	biography	booklet	briefing
broker	charitable	classification	communal	comparatively
compass	concession	configuration	conversely	courtroom
credential	critique	curator	deed	depot
descriptive	desktop	dim	disclosure	documentation
driveway	duo	dynamics	endorsement	enormously
escort	faintly	floral	footage	footstep
franchise	freight	freshly	gala	gamble
garment	gel	genuinely	glimpse	grove

(Level 7)				
hose	incoming	incomplete	instructional	instrumental
interface	interference	interim	lance	lecturer
lengthy	linen	lining	loosely	mattress
mesh	miniature	modem	module	morale
nan	nominee	overwhelmingly	pastry	pavilion
payable	payroll	petition	pier	planner
plaster	plaza	preferably	premature	prestige
promoter	prose	prototype	proximity	readiness
referral	reportedly	retention	rigorous	rip
robust	runway	setup	slate	sponsorship
stocking	striped	successor	superintendent	swap
taxation	thereafter	thickness	tractor	trustee
turf	uneven	unison	unsure	urgency
versatile	vibrant	wardrobe	warrant	

Level 8

furnishing	adjustable	aesthetics	akin	almond
amazement	amend	assortment	auditor	ballroom
barley	blueprint	briefcase	bruise	cartridge
ceramic	clarification	conditional	conveniently	crate
crumb	cuff	culinary	demolition	departmental
dial	disposable	dissatisfaction	distraction	enthusiastically
expressly	forceful	fragrance	freelance	fuzzy
garrison	generously	hairdresser	hen	homemade
hub	imposition	internally	investigative	jersey
landfill	lender	loaf	lumber	metro
multimedia	multinational	occupant	paperback	patronage
pellet	persuasion	plum	porcelain	premiere
presenter	professionalism	progressively	rave	reassurance
rectangular	roster	scanner	seaside	speedy
sponge	spreadsheet	statewide	stroll	suitably
tentatively	thankfully	tier	token	tow
trademark	transitional	trek	trendy	trim
turnout	uniformed	uniquely	universally	unsafe
valuation	walnut			



付録2 TOEIC Bridge L&R に登場し教科書に登場しない語彙

Level 2

additional	appointment	lane	membership	
------------	-------------	------	------------	--

Level 3

aside	avenue	coupon	credit	currently
depart	departure	deposit	electronics	flexible
frank	install	lawn	leadership	maintenance
motor	orientation	reference	regional	retirement

Level 4

appliance	assess	doughnut	expire	fitness
inexpensive	inspector	instruct	laptop	mechanic
participation	rider	technician	update	vegetarian
vice	wireless			

Level 5

aisle	applicant	auto	beverage	builder
caller	catalog	coordinator	dental	password
pedestrian	pharmacy	pickup	purely	renewal
renovate	rubbish	salon	sin	tenant
upcoming				

Level 6

auditorium	cookbook	copyright	elm	holly
paperwork	pat	quartet	ranger	renovation
riverside	roommate	scenic	shoreline	subscribe
subscription	villa			

Level 7

accountant	bedside	booklet	driveway	floral
freshly	pastry	payroll	pier	planner
striped				

Level 8

conveniently	forceful	hairdresser	hen	homemade
loaf	metro	presenter	seaside	trademark
unsafe				

付録3 中高教科書、TOEIC L&R、TOEIC Bridge L&R、共通テストで用いられている高頻度の Ngram

on your answer	your answer sheet	refer to the	to the following
do you know	what does the	mark your answer	answer on your
your answer on	how about you	I have a	look at the
# a.m.	does the speaker	does the woman	look at #
I want to	on to the	to the next	go on to
the next page	does the man	what is the	I don't
the following conversation	you will hear	according to the	a lot of
what will the	it's a	one of the	what do you
the end of	at the graphic	select the best	the best answer
test book and	directions you will	will not be	to do a
you want to	I'd like	would like to	this is the
why does the	do you have	thank you for	I can't
do you want	this is a	do you like	I'm going

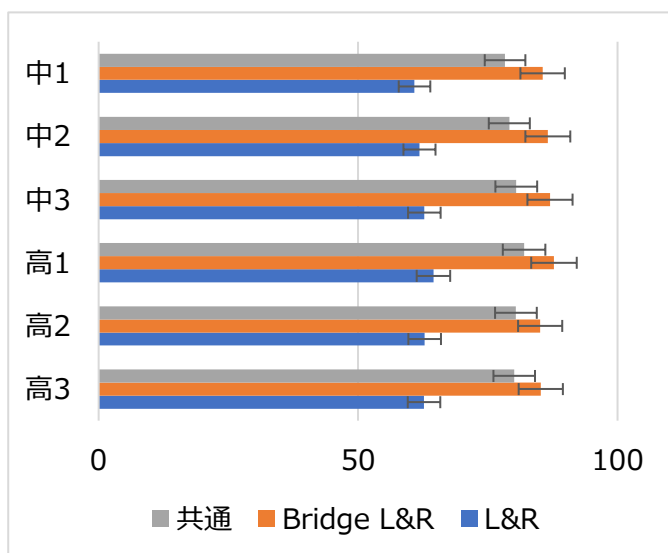
Note. # は数字を表している。

付録4 中高教科書で使用されている語彙が TOEIC L&R、TOEIC Bridge L&R、共通テストで使用されている語彙をカバーする割合

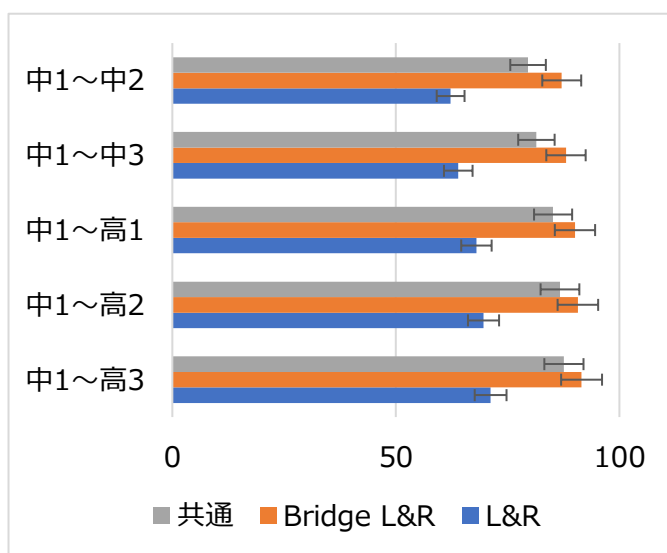
学年	TOEIC L&R	TOEIC Bridge L&R	共通テスト
中1	60.87	85.56	78.32
中2	61.83	86.57	79.14
中3	62.77	86.98	80.48
高1	64.52	87.75	81.99
高2	62.83	85.09	80.42
高3	62.71	85.21	80.09
中1～中2	62.25	87.10	79.57
中1～中3	63.96	88.05	81.43
中1～高1	68.02	90.06	85.17
中1～高2	69.61	90.71	86.71
中1～高3	71.19	91.54	87.59

Note. 計算された数字はいずれもレマ換算したものである。本文でも述べた通り、1人の中高生が複数の教科書を学ぶ可能性は低い点に留意されたい。

教科書の各学年の語彙カバー率



教科書の各学年の累積語彙カバー率

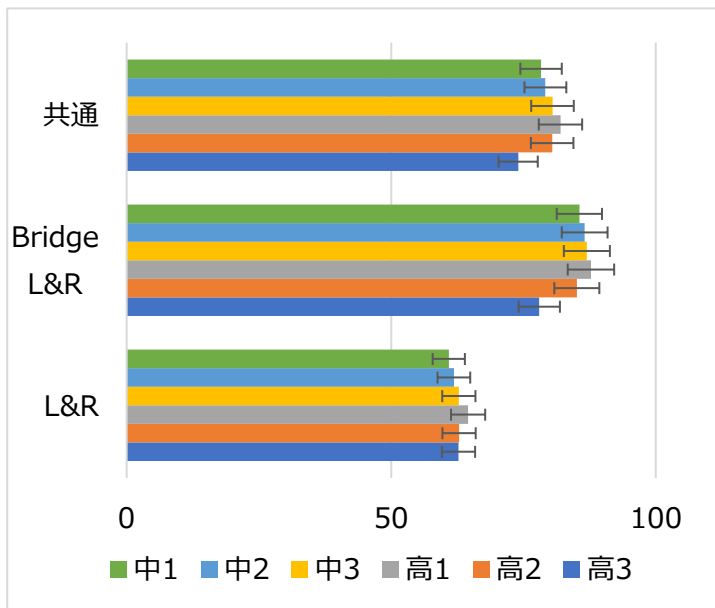


付録5 TOEIC L&amp;R、TOEIC Bridge L&amp;R、共通テストで使用されている語彙が中高教科書で使用されている語彙をカバーする割合

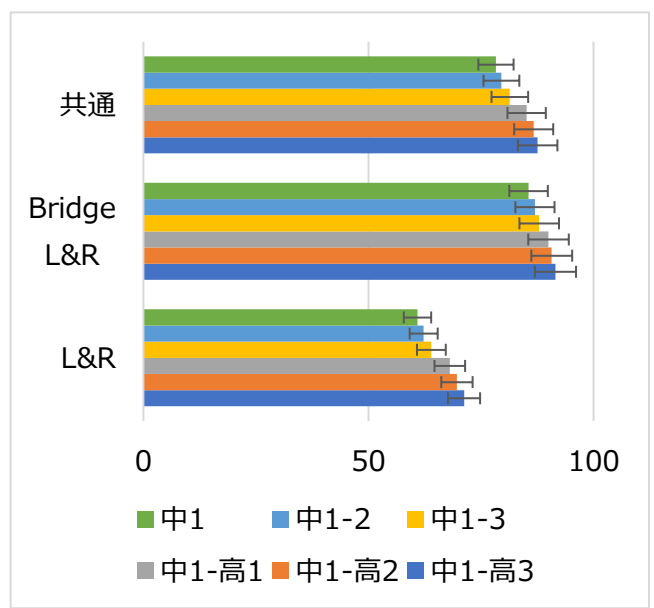
学年	TOEIC L&R	TOEIC Bridge L&R	共通
中1	60.88	85.56	78.32
中2	61.85	86.57	79.14
中3	62.78	86.98	80.48
高1	64.54	87.75	81.99
高2	62.83	85.09	80.42
高3	62.72	77.99	74.00
中1～中2	62.26	86.98	79.54
中1～中3	63.97	87.93	81.40
中1～高1	68.07	90.00	85.13
中1～高2	69.65	90.71	86.71
中1～高3	71.24	91.54	87.59

Note. 計算された数字はいずれもレマ換算したものである。付録4とほとんど同じ数字であった。

各種標準化テストから見た教科書の各学年の語彙カバー率



各種標準化テストから見た教科書の累計語彙カバー率



付録6 個別教科書会社ごとのカバー率一覧

教科書	学年	テスト	カバー率 (%)
A	中学1年	TOEIC L&R	18.53
A	中学1年	TOEIC Bridge L&R	40.00
A	中学1年	共通テスト	27.08
B	中学1年	TOEIC L&R	18.72
B	中学1年	TOEIC Bridge L&R	40.41
B	中学1年	共通テスト	28.00
C	中学1年	TOEIC L&R	16.41
C	中学1年	TOEIC Bridge L&R	37.16
C	中学1年	共通テスト	25.28
D	中学1年	TOEIC L&R	17.45
D	中学1年	TOEIC Bridge L&R	38.28
D	中学1年	共通テスト	25.47
A	中学2年	TOEIC L&R	23.31
A	中学2年	TOEIC Bridge L&R	48.58
A	中学2年	共通テスト	34.38
B	中学2年	TOEIC L&R	24.43
B	中学2年	TOEIC Bridge L&R	52.13
B	中学2年	共通テスト	36.67
C	中学2年	TOEIC L&R	19.24
C	中学2年	TOEIC Bridge L&R	41.48
C	中学2年	共通テスト	28.98

D	中学 2 年	TOEIC L&R	20.72
D	中学 2 年	TOEIC Bridge L&R	43.91
D	中学 2 年	共通テスト	30.48
A	中学 3 年	TOEIC L&R	27.02
A	中学 3 年	TOEIC Bridge L&R	54.32
A	中学 3 年	共通テスト	40.70
B	中学 3 年	TOEIC L&R	27.20
B	中学 3 年	TOEIC Bridge L&R	54.44
B	中学 3 年	共通テスト	40.37
C	中学 3 年	TOEIC L&R	22.62
C	中学 3 年	TOEIC Bridge L&R	46.75
C	中学 3 年	共通テスト	33.99
D	中学 3 年	TOEIC L&R	23.22
D	中学 3 年	TOEIC Bridge L&R	49.41
D	中学 3 年	共通テスト	35.40
A	中学 1 年-2 年	TOEIC L&R	28.54
A	中学 1 年-2 年	TOEIC Bridge L&R	57.22
A	中学 1 年-2 年	共通テスト	41.26
B	中学 1 年-2 年	TOEIC L&R	28.58
B	中学 1 年-2 年	TOEIC Bridge L&R	57.87
B	中学 1 年-2 年	共通テスト	42.34
C	中学 1 年-2 年	TOEIC L&R	24.45
C	中学 1 年-2 年	TOEIC Bridge L&R	51.07
C	中学 1 年-2 年	共通テスト	36.61
D	中学 1 年-2 年	TOEIC L&R	25.18
D	中学 1 年-2 年	TOEIC Bridge L&R	51.24
D	中学 1 年-2 年	共通テスト	36.12
A	中学 1 年-3 年	TOEIC L&R	37.06
A	中学 1 年-3 年	TOEIC Bridge L&R	68.58
A	中学 1 年-3 年	共通テスト	53.41
B	中学 1 年-3 年	TOEIC L&R	35.48
B	中学 1 年-3 年	TOEIC Bridge L&R	66.09
B	中学 1 年-3 年	共通テスト	51.64
C	中学 1 年-3 年	TOEIC L&R	31.35
C	中学 1 年-3 年	TOEIC Bridge L&R	60.24
C	中学 1 年-3 年	共通テスト	46.17
D	中学 1 年-3 年	TOEIC L&R	31.71
D	中学 1 年-3 年	TOEIC Bridge L&R	61.18
D	中学 1 年-3 年	共通テスト	45.61

E	高校 1 年	TOEIC L&R	16.28
E	高校 1 年	TOEIC Bridge L&R	33.49
E	高校 1 年	共通テスト	24.17
F	高校 1 年	TOEIC L&R	21.95
F	高校 1 年	TOEIC Bridge L&R	40.95
F	高校 1 年	共通テスト	33.01
E	高校 2 年	TOEIC L&R	21.95
E	高校 2 年	TOEIC Bridge L&R	45.74
E	高校 2 年	共通テスト	32.91
F	高校 2 年	TOEIC L&R	39.13
F	高校 2 年	TOEIC Bridge L&R	65.15
F	高校 2 年	共通テスト	55.86
E	高校 3 年	TOEIC L&R	23.26
E	高校 3 年	TOEIC Bridge L&R	46.33
E	高校 3 年	共通テスト	34.35
F	高校 3 年	TOEIC L&R	36.61
F	高校 3 年	TOEIC Bridge L&R	61.60
F	高校 3 年	共通テスト	53.21
E	高校 1 年-2 年	TOEIC L&R	27.04
E	高校 1 年-2 年	TOEIC Bridge L&R	52.60
E	高校 1 年-2 年	共通テスト	39.69
F	高校 1 年-2 年	TOEIC L&R	44.32
F	高校 1 年-2 年	TOEIC Bridge L&R	70.53
F	高校 1 年-2 年	共通テスト	61.76
E	高校 1 年-3 年	TOEIC L&R	32.94
E	高校 1 年-3 年	TOEIC Bridge L&R	60.06
E	高校 1 年-3 年	共通テスト	46.95
F	高校 1 年-3 年	TOEIC L&R	51.39
F	高校 1 年-3 年	TOEIC Bridge L&R	76.92
F	高校 1 年-3 年	共通テスト	70.04
A + E	中学 1 年-高校 3 年	TOEIC L&R	46.24
A + E	中学 1 年-高校 3 年	TOEIC Bridge L&R	77.51
A + E	中学 1 年-高校 3 年	共通テスト	63.62
A + F	中学 1 年-高校 3 年	TOEIC L&R	58.04
A + F	中学 1 年-高校 3 年	TOEIC Bridge L&R	84.44
A + F	中学 1 年-高校 3 年	共通テスト	77.60

B + E	中学 1 年-高校 3 年	TOEIC L&R	44.90
B + E	中学 1 年-高校 3 年	TOEIC Bridge L&R	75.27
B + E	中学 1 年-高校 3 年	共通テスト	62.61
B + F	中学 1 年-高校 3 年	TOEIC L&R	56.68
B + F	中学 1 年-高校 3 年	TOEIC Bridge L&R	82.66
B + F	中学 1 年-高校 3 年	共通テスト	76.13
C + E	中学 1 年-高校 3 年	TOEIC L&R	43.21
C + E	中学 1 年-高校 3 年	TOEIC Bridge L&R	73.49
C + E	中学 1 年-高校 3 年	共通テスト	60.35
C + F	中学 1 年-高校 3 年	TOEIC L&R	56.14
C + F	中学 1 年-高校 3 年	TOEIC Bridge L&R	82.07
C + F	中学 1 年-高校 3 年	共通テスト	75.18
D + E	中学 1 年-高校 3 年	TOEIC L&R	43.61
D + E	中学 1 年-高校 3 年	TOEIC Bridge L&R	74.50
D + E	中学 1 年-高校 3 年	共通テスト	60.05
D + F	中学 1 年-高校 3 年	TOEIC L&R	56.70
D + F	中学 1 年-高校 3 年	TOEIC Bridge L&R	83.25
D + F	中学 1 年-高校 3 年	共通テスト	75.70

Note. 中学校教科書は A、B、C、D の 4 社分、高校教科書は E、F の 2 社分である。

付録 7 TOEIC L&R および TOEIC Bridge L&R で使用されている語彙が共通テストで使用されている語彙をカバーする割合

テスト	サブコーパス	共通テスト	
		Listening	Reading
TOEIC L&R	Listening	76.83	63.15
	Reading	88.35	80.89
TOEIC Bridge L&R	Listening	61.38	46.17
	Reading	72.90	52.34

Note. 表中の数値は TOEIC L&R もしくは TOEIC Bridge L&R のリスニング・リーディングセクションで使用されている語彙を全て理解した上で、共通テストで使用されている語彙をどの程度理解できるかを表している。例えば、TOEIC L&R のリスニングセクションで用いられる語彙を全て理解していれば、共通テストのリスニングの語彙の 76.83%、リーディングの語彙の 63.15%を理解できる。

ETS, the ETS logo, PROPELL, TOEIC and TOEIC BRIDGE are registered trademarks of ETS, Princeton, New Jersey, USA, and used in Japan under license.  
Portions are copyrighted by ETS and used with permission.

日本語版発行日：2023年4月

日本語発行：一般財団法人 国際ビジネスコミュニケーション協会

(The Institute for International Business Communication; IIBC)

〒100-0014 東京都千代田区永田町 2-14-2

TEL:050-1790-7396

公式サイト <https://www.iibc-global.org>