

令和 2 年 6 月 22 日現在

機関番号：32510

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K03018

研究課題名(和文) 英語4技能試験に関する研究 TOEIC L&amp;RとS&amp;Wを用いて

研究課題名(英文) A Comparative Study on Listening, Reading, Speaking and Writing Tests Using TOEIC L&amp;R and S&amp;W

研究代表者

神崎 正哉 (Kanzaki, Masaya)

神田外語大学・外国語学部・准教授

研究者番号：30647847

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：本研究ではTOEICの4技能試験の技能別スコア間の関係と各試験に対する受験者の反応を調査した。3年間に渡り、260名の学生からデータを集めた。さらに最終年度に参加した84名には、IELTSも受けてもらい、両試験の比較も行った。

TOEICの技能別スコア間には有意な相関があった。LRスコアと4技能の総合点(SWの合計点に990/400をかけて割合を調整)の間には強い相関があった( $r = .95$ )。これはLRスコアが4技能の総合点の代替になることを意味する。TOEICとIELTSの総合点の間にも強い相関があったが( $r = .79$ )、こちらは代替が可能なレベルではない。

研究成果の学術的意義や社会的意義

本研究では、TOEICの技能別スコア間の関係を調べ、受験者からTOEIC4技能試験に対する意見・感想を尋ねた。英語試験に関するこのような基本情報は、大学入試における英語4技能試験の活用の是非を論じる際に有用である。また、異なる試験の結果を同一の尺度に並べる方法として、CEFRレベルと対照させることが一般的だが、異なる試験間の対照関係を検証する研究は十分に行われていない。本研究では、TOEICとIELTSのスコアを用いて、CEFRレベルとの対照関係の検証を行った。

研究成果の概要(英文)：This study investigated the relationships between the four skill scores of the TOEIC and the test-takers' reactions to each test, or section, of the TOEIC. The study was conducted over a 3-year period, during which data were collected from 260 university students. In addition, 84 participants in the final year of the study took IELTS as well as the TOEIC, and these test results were compared. The four skill scores of the TOEIC significantly correlated with each other, with the listening and reading scores having the highest correlation,  $r = .73$ , and the listening and writing scores having the lowest,  $r = .52$ . The combined listening and reading scores strongly correlated with the overall scores,  $r = .95$ , which suggests that TOEIC LR scores can be a good proxy for overall scores of the TOEIC. In addition, the overall scores of the TOEIC and IELTS strongly correlated with each other,  $r = .79$ .

研究分野：応用言語学

キーワード：TOEIC 英語4技能試験 リスニング リーディング スピーキング ライティング 相関 IELTS

## 様式 C-19、F-19-1、Z-19（共通）

### 1. 研究開始当初の背景

大学入試における英語試験は、リーディングのみ、またはリーディングとリスニングの2技能の能力を試す形で行われるのが一般的であるが、スピーキングとライティングを加えた4技能型試験の導入が推進されるようになって久しい。本研究の開始当初は、4技能試験導入の気運が高まっており、2020年度から始まる共通テストの枠組みに民間の英語4技能試験が組み込まれることになっていた。この大学入試英語成績提供システムは、導入が延期されることになったが、もし予定通り導入されていたら共通テストの受験者約50万人に影響を与える大規模な制度変更になっていた。それにも関わらず、日本人の英語学習者を対象とした英語4技能試験に関する学術的研究は十分に行われていなかった。そのような状況を背景として、本研究は、4技能試験に関する基本的なデータの収集および分析を行うことで英語4技能試験の導入の是非に関する議論に貢献することを目指して始まった。

### 2. 研究の目的

(1) TOEICには、TOEIC Listening and Reading Test（以下 TOEIC LR）と TOEIC Speaking and Writing Tests（以下 TOEIC SW）がある。前者はマークシート方式の紙版テスト、後者はコンピューター版テストである。本研究の主たる目的は、TOEICのリスニング、リーディング、スピーキング、ライティング（以下それぞれ L、R、S、W）の技能別のスコアの間にはどのような関係があるのか、また受験者はそれぞれの試験に対してどのような意見・感想を持つのか調査することであった。

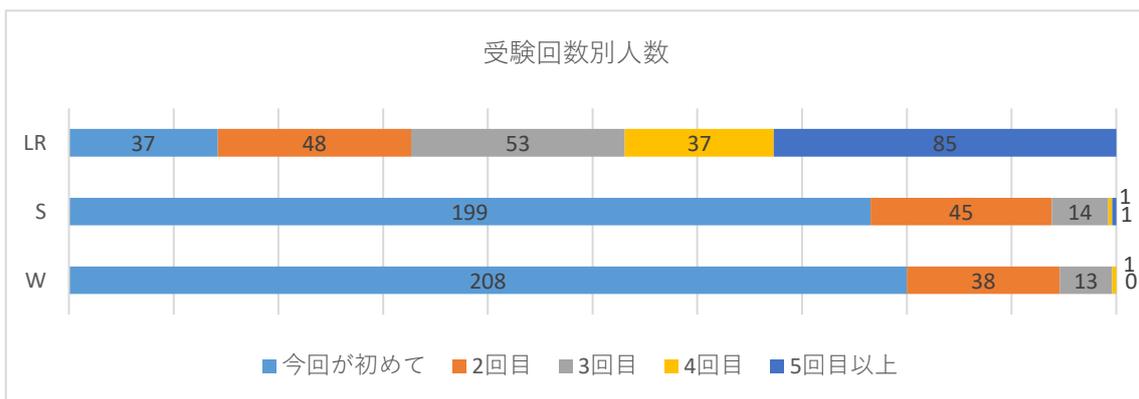
(2) 職場における英語力を測定するための TOEIC と英語圏の大学で授業を受けるのに必要な英語力を測定するための IELTS Academic は、目的が異なる英語試験である。本研究の従たる目的は、TOEIC の4技能試験と IELTS Academic を比較することであった。技能別スコアの比較、総合スコアの比較、CEFR 対照表で同じレベルになるか、コンピューター版テストの TOEIC SW と紙版テストの IELTS SW に対する受験者の意見を調べた。

### 3. 研究の方法

神田外語大学の学生から被験者を募り、TOEIC の4技能試験を受けてもらい、受験後、各テストに対する意見・感想を尋ねるアンケートに答えてもらった。最終年度の参加者には、IELTS も受けてもらった。そして、テストスコアおよびアンケート結果の分析を行った。

### 4. 研究成果

(1) TOEIC の4技能試験に関する研究は、2017年～2019年の3年間に渡り、260名の学生の協力を得て行った。260名の参加者の年度別内訳は、2017年98名、2018年78名、2019年84名、学年別内訳は、1年生73名、2年生72名、3年生71名、4年生44名、性別は女性201名、男性59名であった。専攻別内訳は、英米語102名、国際コミュニケーション96名、国際ビジネスキャリア22名、スペイン語14名、韓国語7名、ポルトガル語7名、中国語5名、インドネシア語7名、ベトナム語2名であった。研究参加時における被験者の TOEIC LR、S、W の受験回数は以下の通りである。



TOEIC スコアの記述統計は以下の通りである(N=260)。

	得点域	平均	標準偏差	最低	最高	歪度	尖度
L	5-495	377.46	67.70	145	495	-0.62	0.52
R	5-495	308.46	77.48	85	470	-0.29	-0.22
S	0-200	128.42	20.39	70	190	0.02	0.42
W	0-200	140.85	21.45	60	190	-0.70	0.82

LR	10-990	685.92	135.07	270	965	-0.37	-0.01
SW	0-400	269.27	119.63	140	370	-0.38	0.42
総合点*	10-1980	1352.36	209.97	715.75	1855.75	-0.39	0.38

\*総合点は、SW スコアに 990/400 をかけて LR スコアに可算して求めた。

TOEIC の技能別スコア間の相関係数 (Pearson's  $r$ ) は、以下の通りである (All  $ps < .001$ )。

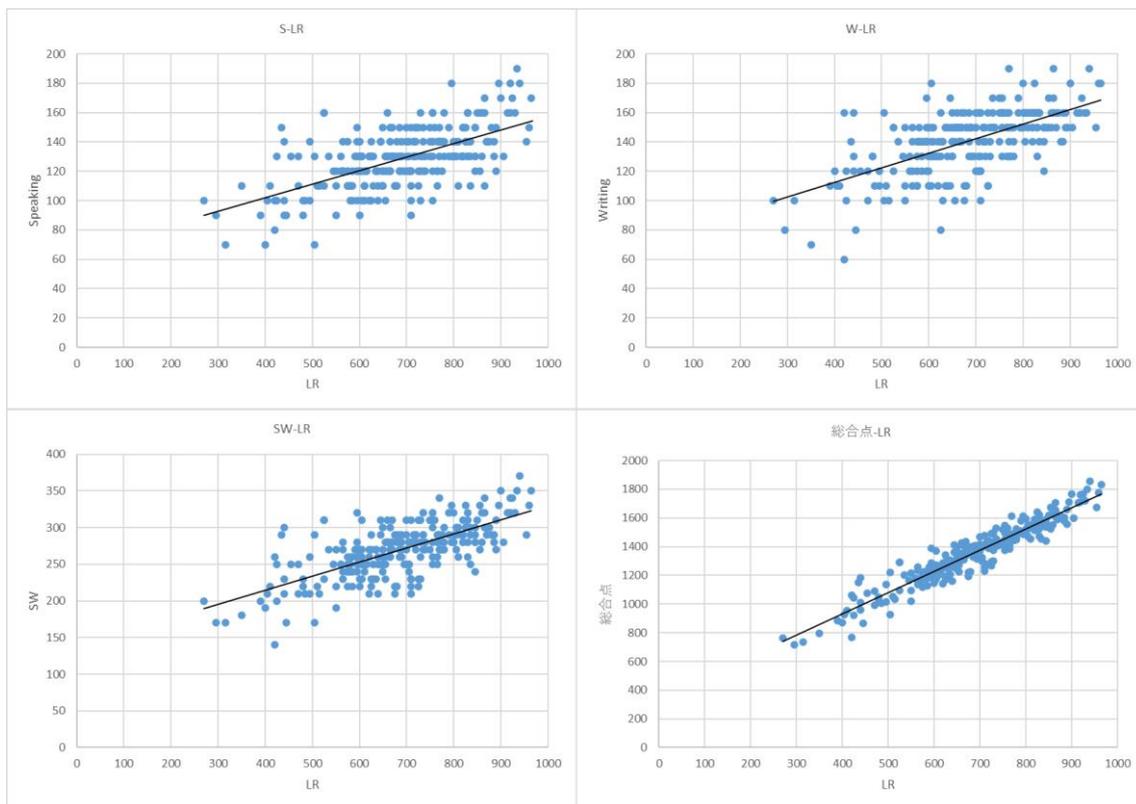
	L	R	S	W
L	1	.73	.55	.52
R		1	.59	.64
S			1	.58
W				1

LR スコアと S スコア、W スコア、SW スコア、総合点 (SW スコアに 990/400 をかけて LR スコアに可算) との間の相関係数 (Pearson's  $r$ ) は以下の通りである (All  $ps < .001$ )。

	S	W	SW	総合点
LR	.62	.63	.70	.95

Dorans (2004) は、2 つのテスト間に .866 以上の相関があれば、一方のスコアから他方のスコアの妥当な予測が可能であることを示している。この基準に従うと、LR スコアは S スコア、W スコア、SW スコアの妥当な代替とはならないが、総合点の妥当な代替となり得る。

LR スコアと S スコア (左上)、W スコア (右上)、SW スコア (左下)、総合点 (右下) の散布図を以下に示す。



総合点と技能別スコア間の相関係数 (Pearson's  $r$ ) は以下の通りである (All  $ps < .001$ )。

	L	R	S	W
総合点	.86	.91	.78	.80

総合点との間の相関が一番高いのは R スコアで、 $r = .91$  は、総合点の妥当な代替となり得るレベルである (Dorans, 2004)。

上記の結果は、LR スコアまたは R スコア単独で総合点の代替となり得ることを示している。

ここから 4 技能の総合力を予想するのに SW テストは不要であるという議論が成り立つ。ただし、本研究の参加者は外国語学部の学生で、平均的な日本人英語学習者よりスピーキングとライティングの練習を多く積んでいる点を考慮する必要がある。スピーキングやライティングの練習量の少ない学習者が TOEIC の 4 技能試験を受けた場合、異なる結果が出る可能性もある。

アンケートの結果から、受験時の緊張度合いは  $S > W > L > R$ 、問題の難易度は  $S > W > R > L$ 、楽しさレベルは  $S > W > L > R$ 、好きかどうかは  $L > S > W > R$  の順であることが分かった。S と W は試験の問題形式に不慣れなため、自分の実力を発揮できなかったと感じる参加者が多くいたが、SW ともそれぞれスピーキング力・ライティング力を測定するのに良いテストだという回答が多かった。S に関してはコンピューター版テストより人と話す面接式テストの方が自分の実力を出せると答えた参加者が多かったが、W に関しては紙版テストよりコンピューター版テストの方が自分の実力を出せるという回答が多かった。

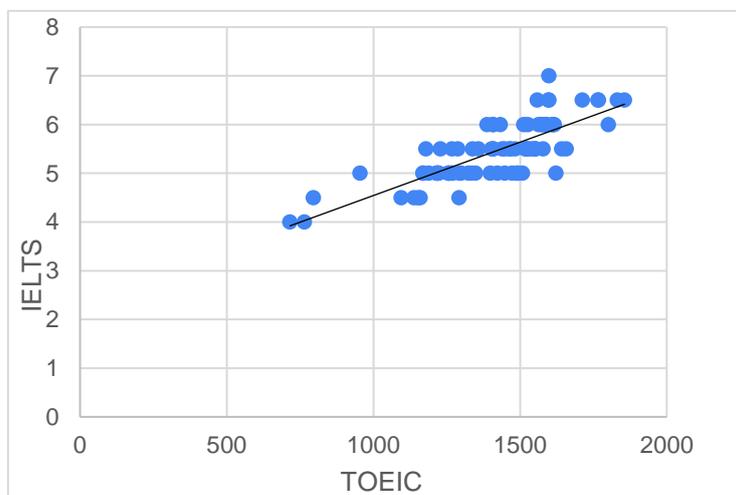
(2) 2019 年の参加者 84 名には IELTS Academic も受けてもらい、TOEIC と IELTS の比較を行った。スコアの記述統計は以下の通りである ( $N = 84$ )。T は TOEIC を、I は IELTS を意味する。

	得点域	平均	標準偏差	最低	最高	歪度	尖度
T-L	5-495	383.15	76.28	145	495	-0.89	0.78
T-R	5-495	330.06	80.19	85	470	-0.82	0.69
T-S	0-200	137.74	19.35	90	190	0.21	0.34
T-W	0-200	144.76	22.41	70	190	-0.92	1.88
T-総合点*	10-1980	1412.40	221.58	715.75	1855.75	-0.76	1.21
I-L	0-9	5.30	0.68	3.5	7.0	0.10	0.35
I-R	0-9	5.46	0.93	3.5	8.0	0.30	-0.09
I-S	0-9	5.54	0.82	3.5	8.0	0.20	0.26
I-W	0-9	5.33	0.51	3.5	6.0	-0.67	0.68
I-総合点	0-9	5.45	0.61	4.0	7.0	0.13	-0.16

\*TOEIC の総合点は、SW スコアに 990/400 をかけて LR スコアに可算して求めた。

TOEIC と IELTS の技能別スコア間の相関は、L 間  $r = .64$ 、R 間  $r = .69$ 、S 間  $r = .50$ 、W 間  $r = .58$ 、総合点間の相関は  $r = .79$  であった (All  $ps < .001$ )。総合点間の相関は、技能別スコア間の相関より高くなっているが、妥当な代替となり得る最低値の .866 (Dorans, 2004) を下回っている。これは、TOEIC と IELTS は異なる能力を測定しており、両テスト間の代替は妥当ではないことを意味する。よって、両テストを同じ換算表上に並べて対照させるのは好ましいことではない。

TOEIC と IELTS の総合点の散布図を以下に示す。



文部科学省が 2018 年 3 月に発表した「各資格・検定試験と CEFR との対照表」では、TOEIC と IELTS の CEFR レベルは以下ようになっていた (TOEIC は SW スコアを 2.5 倍して LR スコアと合算)。

CEFR	A2	B1	B2	C1	B1
TOEIC	625-1145	1150-1555	1560-1840	1845-1990	1150-1555
IELTS	該当なし	4.0-5.0	5.5-6.5	7.0-8.0	4.0-5.0

この基準に則って、84名の参加者のTOEICとIELTSの両方のCEFRレベルを決めると以下のような人数の内訳になる。

	IELTS B1	IELTS B2	IELTS C1
TOEIC C1	0	1	0
TOEIC B2	1	22	1
TOEIC B1	27	26	0
TOEIC A2	6	0	0

84名のうち、TOEICとIELTSのCEFRレベルが同じになった人が49名、IELTSの方が1レベル高く出た人が33名、TOEICの方が1レベル高く出た人が2名いた。この結果からIELTSを受けた方がTOEICを受けるよりCEFR換算でよい結果が出やすいと言える。

両テストのスピーキングに関しては、面接官と話す方式のIELTSの方が自分の実力を発揮できると答えた人が56名、コンピューター版のTOEICの方が自分の実力を発揮できると答えた人が15名、どちらも変わらないと答えた人が13名いた。IELTSを好む理由としては、人と話す方がコンピューターに向かって話すよりも話しやすいという意見が多かった。TOEICを好む理由としては、人と話す緊張する、面接官が怖かった・やる気がなかったなどが挙げられていた。

ライティングに関しては、紙版のIELTSの方が自分の実力を発揮できると答えた人が38名、コンピューター版のTOEICの方が自分の実力を発揮できると答えた人が31名、どちらも変わらないと答えた人が15名であった。キーボード入力に慣れているかどうかで、紙版またはコンピューター版の好みが決まる傾向があった。

#### <引用文献>

Dorans, N. J. (2004). Equating, concordance, and expectation. *Applied Psychological Measurement*, 28(4), 227-246. <https://doi.org/10.1177/0146621604265031>

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 1件/うちオープンアクセス 1件）

1. 著者名 Masaya Kanzaki	4. 巻 Issue 2019
2. 論文標題 TOEIC Listening and Reading Test and Overall English Ability	5. 発行年 2020年
3. 雑誌名 JALT Postconference Publication	6. 最初と最後の頁 未定
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計2件（うち招待講演 0件/うち国際学会 2件）

1. 発表者名 Masaya Kanzaki
2. 発表標題 Test-Taker Reactions to TOEIC L&R and S&W Tests
3. 学会等名 JALT2019（国際学会）
4. 発表年 2019年

1. 発表者名 Masaya Kanzaki
2. 発表標題 IELTS vs. TOEIC: Scores and Test-Taker Reactions
3. 学会等名 JALT2020（国際学会）
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

英語4技能試験に関する研究 TOEIC L&RとS&Wを用いて (researchmap 研究ブログ)  
[https://researchmap.jp/blogs/blog\\_entries/view/98804/35349e134c8a00361c3e129d212387da?frame\\_id=543872](https://researchmap.jp/blogs/blog_entries/view/98804/35349e134c8a00361c3e129d212387da?frame_id=543872)  
 researchmapの研究ブログを介して本研究で集めた元データを公開。

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	澁谷 由紀  (Shibuya Yuki)  (80648155)	神田外語大学・外国語学部・准教授   (32510)	
研究分担者	上原 雅子  (Uehara Masako)  (00787434)	神田外語大学・外国語学部・講師   (32510)	