

TOEIC Bridge テストと TOEIC テストを用いた英語力測定： 到達目標型教育における確かな英語力測定を目指して

前 田 啓 朗

広島大学外国語教育研究センター

1. 研究の背景

教育には評価を欠かすことができないため、英語教育においても、学習者の英語力を測定する場面がある。測定に用いる方法としては、いわゆるペーパーテストや言語活動中の観察、面接などの複数の種類が用いられるが、特に指導と密接に関連して行われる評価の場合には、指導過程において適した場面、適した方法で、複数回の測定を実施することが望ましいとされる。一方、ある程度の期間の指導について、複数のクラスを横断的に総括して評価するためには、個別のクラスに特有の言語材料や言語機能にとらわれない、一般的な英語コミュニケーション能力を測定することが必要となる。

広島大学では、目標を策定し、目標に沿って指導を行い、目標を到達できたかどうかを点検し、必要であれば教育課程等の見直しを図るというシステムに沿い、これを到達目標型教育と称して制度改革を進めている。その一環として、英語教育においては日常的、一般的なコミュニケーション能力の尺度として TOEIC テスト (Educational Testing Service, 2004a) を用いている。試験日時を自由に設定できる団体受験特別制度 (IP: Institutional Program) を利用して、平成15年度入学生より年次進行で、1年次5月、1年次2月、2年次2月、3年次2月の合計4回、学生に受験を求めている。

ただし、そのスコア分布 (広島大学外国語教育研究センター, n.d.) を見ると、入学時に合計スコアが400点を下回る学生が約半数弱いることがうかがえる。TOEIC テストは非常に英語能力の高い学習者から低い学習者までを測定可能であるために、問題の難易度も幅広くならざるをえない。その結果、総項目数は200問で総実施時間は120分という構成になっている。英語力の低い学習者にとっては理解可能な項目が少ない中で集中力を持続させることが困難であることが予想され、英語力を正確に測定できなくなる危険性があるといえる。

したがって、本研究では、同様の目的で作成された TOEIC Bridge テスト (Educational Testing Service, 2004b) に着目する。平成13年度から開始されたこのテストは、TOEIC のおおむね450点以下の学習者を目安として想定しており、それ以上のスコアを取る学習者には TOEIC を勧めている (国際ビジネスコミュニケーション協会, n.d.)。また、総項目数は100問で総実施時間は60分であり、TOEIC テストの半分である。このため、英語力が低い学習者にとっても受験の負担が少ないであろうと推測できる。

このことから測定の利便性だけを考えると、広島大学の半数強の学生の英語力を測定するためには、TOEIC テストを用いるまでもなく TOEIC Bridge テストで十分であるとも考えられる。したがって、広島大学の学生の英語力を測定するためには TOEIC Bridge テストによる測定がどの程度有効に機能するのかということを検討するために、測定の安定性や練習効果について検証を試みるとともに、TOEIC Bridge テストのスコアと TOEIC テストのスコアの関連についても分析を行う。

2. 調査

2.1 被調査者

広島大学の1年次生(平成16年度入学生)を対象として、参加者を募集した。参加申し込みの段階で、研究目的でデータを利用することや分析結果の公開時に個人情報は保護することなどについて文書で説明し、同意を得た。156名が参加の申し込みを行い、これをランダムに2群に割り当てた。また、広島大学が行うTOEIC IPテスト全学一斉実施のデータに欠損があるケースを削除し、最終的に141名分のデータを分析に用いた。なお、被調査者は全員、この調査以前にTOEIC BridgeテストおよびTOEICテストの受験経験がなかった。

2.2 材料

TOEIC Bridgeテストの団体受験特別制度を用いて、再検査信頼性も検討するために2種類のフォームを用いることとした。調査の対象である学習者集団には多様な英語力を持つ学生がいるために天井効果も危惧されたが、TOEICテストよりはかなり簡単であることに、参加者を募集する段階で言及しておけば問題ないと判断した。なお、スコアは2点刻みであり、リスニング・セクションおよび、リーディング・セクションがともに90点満点で、合計180点が最高スコアである。

TOEICテストについては、広島大学が行うTOEIC IPテスト全学一斉実施のデータを用いた。対象となる学生は5月と2月に異なるフォームでこのテストを受験している。スコアは5点刻みで、リスニング・セクションおよび、リーディング・セクションがともに495点満点の合計990点が最高スコアである。

2.3 方法

TOEIC Bridgeテストの再検査信頼性と練習効果についても検討するため、被調査者を無作為にA群とB群の2群に分割した。TOEIC Bridgeテストのフォームは2種類(フォーム1・フォーム2)であり、A群($n=68$)にはフォーム1・休憩等・フォーム2、B群($n=73$)にはフォーム2・休憩等・フォーム1の順で、平成16年5月8日および9日に実施した。

TOEICテストのスコアには、5月分として平成16年5月15日および23日に実施されたTOEIC IPテスト全学一斉実施のデータ、2月分として平成17年2月11日および21日に実施されたデータを用いた。

これらのデータについて、相関分析を行うとともに、TOEICテストのスコアを従属変数とした回帰分析によって、TOEIC Bridgeテストのスコアとの関連を分析した。

2.4 結果と解釈

2.4.1 TOEIC Bridge

被調査者全員のスコアは表1に示される。A群についてはフォーム2、B群についてはフォーム1と、2群のいずれにおいても2回目に受験したフォームによるスコアの方が平均して高い。ただし、スコアは2点刻みであり、2回目に受験した方が合計スコアにおいてA群では2.2点、B群では2.1点高いという差は、実質的に意味がある差であるとは解釈できない。

また、標準偏差についても2つの群および2つのフォームについて同様の値が得られている。これらのことから、TOEIC Bridgeテストの受験回数による練習効果は無視できる程度のものであり、次に検討する相関係数とあわせて、再検査の信頼性は十分であると判断できる。

表1 TOEIC Bridge スコア

		フォーム1		フォーム2		備 考
		平均	標準偏差	平均	標準偏差	
A 群 (n=68)	合計	153.7	10.5	155.9	11.2	フォーム1, 2の順 に受験
	リスニング	73.8	7.4	74.7	7.9	
	リーディング	79.9	5.4	81.2	4.9	
B 群 (n=73)	合計	156.8	10.5	154.7	11.3	フォーム2, 1の順 に受験
	リスニング	76.6	6.4	74.0	7.2	
	リーディング	80.2	6.0	80.7	5.8	
全体 (N=141)	合計	155.3	10.6	155.3	11.2	
	リスニング	75.3	7.0	74.4	7.5	
	リーディング	80.1	5.7	80.9	5.4	

受験順序による練習効果はないものと考え、2つのフォーム間の相関について検討する。両フォームにおけるセクションごとのスコアと合計スコアという6つの変数について、ピアソンの積率相関係数を算出した(表2)。フォーム間の相関は合計スコアにおいて $r=.69$ であり、セクション別にはさらに低い。両フォームが別の構成概念を測定するものであるというより、単なる回帰効果であると判断するのが妥当であろう。また、同一フォーム内のリスニング・セクションとリーディング・セクションの相関係数について、フォーム1では $r=.38$ 、フォーム2では $r=.51$ であることから、2つのセクションは弁別可能な能力を評定していると判断できる。

表2 TOEIC Bridge スコアにおけるセクション別の相関係数

		フォーム1			フォーム2		
		合計	リスニング	リーディング	合計	リスニング	リーディング
フォーム1	合計	1.00	.87	.79	.69	.64	.56
	リスニング	.87	1.00	.38	.61	.64	.38
	リーディング	.79	.38	1.00	.54	.40	.56
フォーム2	合計	.69	.61	.54	1.00	.91	.82
	リスニング	.64	.64	.40	.91	1.00	.51
	リーディング	.56	.38	.56	.82	.51	1.00

N=141

2.4.2 TOEIC

表3に、平成16年5月と平成17年2月におけるTOEICテストのスコアを示す。先述のように、TOEIC BridgeテストはTOEICテストのスコア450点以下の学習者を目安として対象にしている。今回の被調査者はややその目安よりも英語力の高い学生であったことがわかる。このことは、表1に示されたTOEIC Bridgeスコアの代表値が上限に近いことからうかがえる。

また、合計スコアにおける5月と2月のピアソンの積率相関係数は $r=.72$ であり(表4)、2回のTOEIC Bridgeテストと同様に回帰効果が起きていると思われる。換言すれば、全体的に英語力が向上しており、あまり入れ替わりが起きていると考えられる。また、TOEIC Bridgeテストと同様に、2つのセクションは十分に弁別されていると解釈できる。

表3 TOEIC スコア

	5月		2月	
	平均	標準偏差	平均	標準偏差
合計	511.4	105.9	535.7	120.2
リスニング	276.5	59.4	286.4	69.6
リーディング	234.9	59.0	249.4	62.1

N=141

表4 TOEIC スコアにおけるセクション別の相関係数

		5月			2月		
		合計	リスニング	リーディング	合計	リスニング	リーディング
5月	合計	1.00	.90	.89	.72	.69	.62
	リスニング	.90	1.00	.60	.67	.72	.48
	リーディング	.89	.60	1.00	.63	.52	.63
2月	合計	.72	.67	.63	1.00	.92	.90
	リスニング	.69	.72	.52	.92	1.00	.67
	リーディング	.62	.48	.63	.90	.67	1.00

N=141

2.4.3 TOEIC と TOEIC Bridge

5月のTOEICスコアとTOEIC Bridgeスコアについて、合計スコアの分析を行い、その後にセクションごとの分析を行う。これらのスコアは同じ月に実施されたテストによるものであるため、特にTOEICスコアと2種のTOEIC Bridgeスコアとの関連を検討し、予測を試みる。

2.4.3.1 5月のTOEIC合計スコアと2種のTOEIC Bridge合計スコア

合計点の散布図を、図1、図2に示す。図1の2変数間の相関は $r=.65$ であり、図2においては $r=.75$ であった。フォームによってTOEICスコアとの相関が若干異なるものの、図および係数から判断して、これらは十分に意味のある共変関係にあり、一方でもう一方をある程度予測できると考えられる。

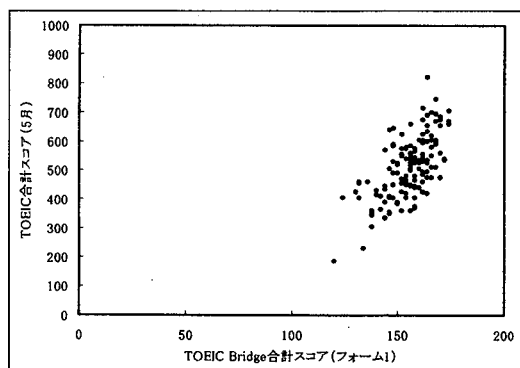


図1 TOEIC (5月) と TOEIC Bridge (フォーム1) における合計スコア (N=141)

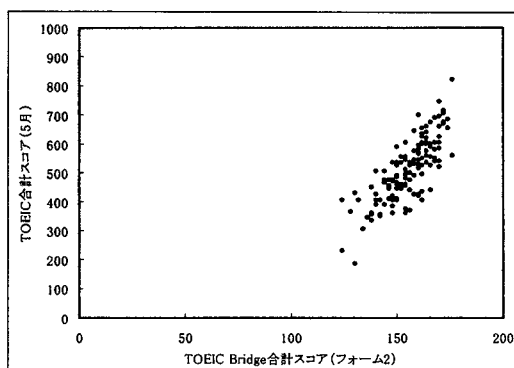


図2 TOEIC (5月) と TOEIC Bridge (フォーム2) における合計スコア (N=141)

表5 2種のTOEIC Bridge 合計スコアによるTOEIC 合計スコア（5月）の回帰分析

	独立変数：TOEIC Bridge 合計スコア（フォーム1）		独立変数：TOEIC Bridge 合計スコア（フォーム2）	
	係数	標準誤差	係数	標準誤差
回帰係数	6.5	0.6	7.4	0.5
切片	-491.2	100.8	-642.8	76.4

N=141, 従属変数：TOEIC 合計スコア（5月）

回帰分析を行った結果、表5のような回帰係数と切片が得られた。これらのテストはほぼ同時期に行われているため、TOEIC Bridge の合計スコアでTOEIC 合計スコアを予測するには、次の式のいずれかを用いることができる。なお、係数や項を整理して、TOEIC 合計スコアによってTOEIC Bridge スコアを推定することもできる。

TOEIC 合計スコア

$$= 6.5 \times \text{TOEIC Bridge 合計スコア} - 491.2 \quad (\text{フォーム1の係数から})$$

TOEIC 合計スコア

$$= 7.4 \times \text{TOEIC Bridge 合計スコア} - 642.8 \quad (\text{フォーム2の係数から})$$

ただし、切片だけに着目しても100.8や76.4と、標準誤差も大きい。上の式はあくまで、本研究におけるテスト受験者を対象にした場合の簡便な推定法であり、一般化には慎重になるべきであろう。

2.4.3.1 5月のTOEIC セクション別スコアと2種のTOEIC Bridge セクション別スコア

表2および表4に示されたように、リスニング・セクションとリーディング・セクションの相関は、両者が別の能力を評定していると判断できる程度に低い。そこで、TOEIC スコアとTOEIC Bridge スコアについて、セクション別に検討を行う。

リスニングスコアに関して、図1および図2と同様に、散布図を示す（図3、図4）。相関係数は図3において $r=.59$ 、図4において $r=.77$ であった。

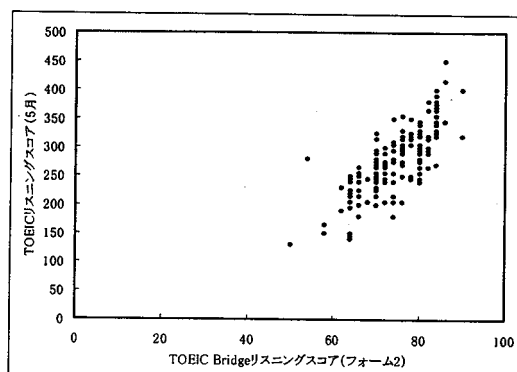
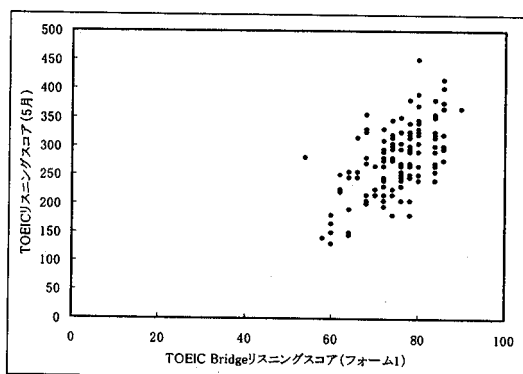


図3 TOEIC（5月）とTOEIC Bridge（フォーム1）におけるリスニングスコア（N=141）

図4 TOEIC（5月）とTOEIC Bridge（フォーム2）におけるリスニングスコア（N=141）

同様に、リーディングスコアについても図5、図6に散布の状況を示す。相関係数はそれぞれ $r=.55$ および $r=.65$ であった。

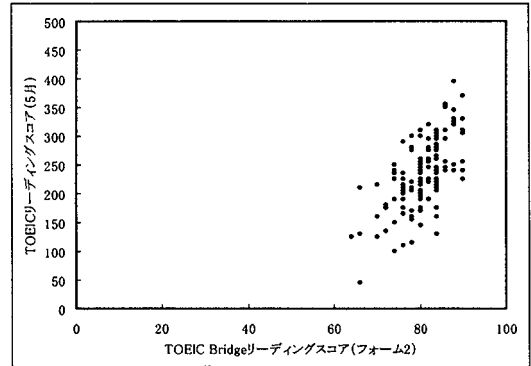
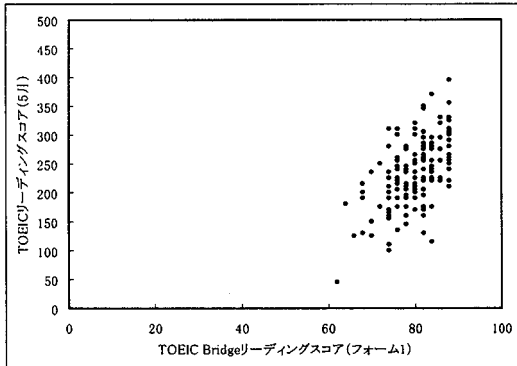


図5 TOEIC(5月)とTOEIC Bridge(フォーム1)におけるリーディングスコア (N=141) 図6 TOEIC(5月)とTOEIC Bridge(フォーム2)におけるリーディングスコア (N=141)

両セクションに共通して、TOEIC Bridge フォーム1のスコアの方が、TOEICスコアとの相関が低い。リーディングスコアについては、散布図から天井効果が起きていることがわかる。先述のように、TOEIC Bridge テストはTOEICテストの合計スコアが450点以下の受験者を目安としているため、今回の受験者集団にとっては易しすぎたと言える。リスニングスコアに関しては、図3の散布図は図4の散布図に比べて、外れ値として考えられるケースの数は同様であるが、全体的に直線的な図4に比べて少し膨らんだ様子が見て取れる。テストの性質上、項目や個々の回答状況を検討することはできないので、この原因の検討は本研究の限界を超える。

ただし、その他においては特に問題がないことから、回帰効果のために相関係数が控えめになっていることを考慮すると、TOEICスコアとTOEIC Bridgeスコアにはある程度の直線的な関係があると考えられる。

3. 結論

被調査者の英語力がTOEIC Bridge テストに適する英語力よりは全体的に高めであったことから、データが天井効果を示すなどの問題があった。しかし、TOEIC Bridge テストの目安とされるのはTOEICスコア450点以下である一方で、600点、700点レベルの受験者にとってもTOEIC BridgeスコアとTOEICスコアはある程度直線的な関係を持ち、高レベルの学習者の学力を測定することにも対応できるという可能性が示唆された。

国際ビジネスコミュニケーション協会(n.d.)によって「TOEICスコア450点に達している方には、TOEICの受験をお勧めします」もしくは「TOEIC Bridgeの受験目安としているTOEICスコア450点以下は、日本独自のガイドラインとしてTOEIC運営委員会より示しています」と提案されている。前者の意図としては、TOEICスコア450点以上の学習者も、ある程度まではTOEIC Bridgeでも測定可能であるが、その程度までの英語力を身に付けた学習者にはより真正性の高いTOEICテストを受験する方が望ましいことが考えられる。後者の意図は、文字通り、学習者の英語力に即したテストを受験するほうが望ましいということであろう。

本研究の被調査者は、5月の段階でTOEICスコアのセクション別平均がリスニングで276.5点、リーディングで234.9点である。このスコアを、*Can-Do Guide linking TOEIC* (Educational Testing Service, 2000)で見ると、リスニングでは何一つ“Can Do”に該当するものはなく、“understand simple questions in social situations such as ‘How are you?’ ‘Where do you live?’ and ‘How do you feel?’”という項目でさえも“Can Do with Difficulty”に含まれる。リーディングでは“read, on storefronts, the type of store or services provided (e.g., ‘dry cleaning,’ ‘book store’)”だけが“Can Do”に含まれる。

広島大学における英語教育の目標を考慮すると、当然ながらこれよりもさらに高度な英語の運用能力を到達目標とするものである。したがって、TOEIC Bridgeテストによる測定の方が、TOEICテストを用いるよりも受験者に対する負担が少ないことなどの利点が挙げられるが、広島大学の教育目標という観点からは、TOEICテストを用いることによって目標と合致した測定を行う方が望ましいと考えられる。TOEIC Bridgeテストについては、英語力が低い一部の学習者に対する補完的な測定等に用いるという用途が考えられる。

なお、今回の調査では調査参加者の英語力が予期したものよりも高くTOEIC Bridgeテストで天井効果が起こったこと、調査参加者の背景が広島大学の1年次生だけに偏っていることなどが限界であり、TOEIC BridgeテストおよびTOEICテストに関する知見を安易に般化することはできない。今後の追調査や複数の研究結果の比較検討が必要になろう。

付記

本研究は、科学研究費補助金 若手(B) 16720130 による研究成果の一部である。

注

TOEICおよびTOEIC Bridgeは、米国ETSの登録商標である。

引用文献

- Educational Testing Service. (2000). *Can-do guide linking TOEIC*. Retrieved December 10, 2004, from http://ftp.ets.org/pub/corp/ell/TOEIC_Can_Do.pdf
- Educational Testing Service. (2004a, May 15). *TOEIC home*. Retrieved December 10, 2004, from <http://www.ets.org/toeic/>
- Educational Testing Service. (2004b, March 13). *TOEIC Bridge home*. Retrieved December 10, 2004, from <http://www.ets.org/toeicbridge/>
- 国際ビジネスコミュニケーション協会. (n.d.). *TOEIC(R) Bridge*. Retrieved December 10, 2004, from <http://www.toeic.or.jp/bridge/>
- 広島大学外国語教育研究センター. (n.d.). 『TOEIC(R) IP 全学一斉実施』 Retrieved December 10, 2004, from <http://home.hiroshima-u.ac.jp/flare/toeicip/>

ABSTRACT

**The TOEIC Bridge Test and TOEIC Test as Measurement Tools:
Toward an Effective Measurement in Criterion-Referenced Evaluation**

Hiroaki MAEDA

Institute for Foreign Language Research and Education,
Hiroshima University

This paper investigates two ways in which the English proficiency of students at Hiroshima University is measured: the TOEIC Bridge Test and the TOEIC Test. By administering two different forms of the TOEIC Bridge Test to participants, its test-retest reliability is confirmed. Correlation analyses between scores on the listening and reading sections show that the two sections reasonably measure different components of English proficiency. TOEIC Bridge Test scores were compared with TOEIC Test scores, using correlation analyses and regression analyses. As a result, it is shown that the two kinds of test scores correlate almost linearly, and that the estimation formulas can be used with caution to convert the two kinds of scores.