

千葉大学におけるTOEIC IPとTOEFL ITPの スコア分析と経年調査

A Descriptive and Longitudinal Analysis of
TOEIC IP and TOEFL ITP Scores at Chiba University

土肥 充, 張 智君

1. はじめに

近年、我が国では官民を挙げて「グローバル人材の育成」が声高に叫ばれ、政策会議としての「グローバル人材育成推進会議」の「審議まとめ」（グローバル人材育成推進会議、2012）では、大学の入試や単位認定、初等中等教育の英語担当教員の採用や研修、国家公務員の採用、等へのTOEICとTOEFLの活用に言及している。このような動きと呼応するかのよう、国内のTOEIC受験者数は2011年に年間200万人を超え、その半数以上が企業や大学などの団体が実施する「団体特別受験制度（IP: Institutional Program）」の受験者である（国際ビジネスコミュニケーション協会、2013a）。一方のTOEFLは正確な受験者数は公表されていないが、1964年以来実施されてきたTOEFL PBT/CBT/iBTの3つの形式のテストを世界規模の累計で2700万人以上が受験した（Educational Testing Service, 2013a）。これとは別に団体向けテストプログラムであるTOEFL ITP（Institutional Testing Program）が正規のTOEFLよりも大幅に低価格で実施されている。

千葉大学では2003年度から外国語センター（現、言語教育センター）が全学の学生を対象にTOEIC IPを実施し、2007年度からはTOEFL ITPを実施してきた。このうちTOEIC IPの結果について、土肥（2006）は2003～2004年度の2年分のべ1,240名の分析をし、土肥・柳瀬（2009）は2003～2007年度の5年分のべ6,689名の分析を行った。後者は前者よりも大規模なデータを収集しただけでなく、より多様な分析方法を採用して、信頼性の高い包括的な研究となった。その後5年が経過し、10年間で本学のTOEIC IPの受験者数は2万名を超えた。また、TOEFL ITPについても6年間で500名を超えて、信頼性の高い分析が可能な規模のデータが蓄積された。さらに、ListeningとReadingのSectionからなるTOEIC IPとは別にTOEIC Speaking/Writing IPテストも開始され、本学では2011年度と2012年度に1回ずつ実施した。本研究では、上記のTOEICの結果分析の二論文を発展させた研究を行うことを目指した。

2. 研究の目的

本研究の主な目的は、1) 千葉大学学生の英語力について大規模な経年調査を行うこと、2) TOEIC IP, TOEFL ITP, TOEIC Speaking/Writing IP Tests のスコアの比較と関連

性の研究を行うことである。

3. 研究の方法

本研究は、筆者が入手し得る過去10年のすべてのデータに基づき、以下の(1)～(4)の方法で行った。全学の外国語教育を統括する立場の言語教育センターが直接関わらない形で、学部・学科等の単位で独自に実施されたテストについては本研究の対象外とした。

(1) TOEIC IP実施と受験者数の概要

TOEIC IPは2003年度以降10年間(2003年5月～2013年2月)にわたって、年4～6回計55回実施し、そのうち2007年度以降は年1回ずつ計6回、受験料大学負担で全学部の1年次学生に受験を義務付け、その他の49回は希望者を対象に有料受験という形態で実施した。この49回の一部では、一部の学部や学科が受験を義務付けた場合も含まれ、詳細については土肥・柳瀬(2009)の表4を参照されたい。また、正規学生以外の科目等履修生、聴講生、教職員等も受験しているが、本研究は正規学生のみを対象とし、それ以外の受験者は分析に一切含めないこととする。合計受験者数は、のべ21,280名(学部生20,439名、大学院生841名)で、うち2,172名(学部生2,082名、大学院生90名)は同一学生が複数回受験した。学部生は前述の通り2007年度以降、1年次で1回の受験を最低限義務付けられているため、今回の全受験者の半数以上は1年次学生であった。なお、同一学生が複数回受験したかどうかの判断は、学生証番号に基づいて一意に決定される受験番号によって識別しているが、学部生が大学院に進学した場合は学生証番号が変わってしまうため、同一学生と扱うことはできなかった。また、2008年度まで大学院生の受験番号は学生証番号と無関係の受験番号を使用していたため、大学院生については2009年度以降で同一学生かどうかの判断を行った。

(2) TOEFL ITP実施と受験者数の概要

TOEFL ITPは2007年度以降6年間(2007年12月～2012年12月)にわたって、年1～2回計11回実施し、いずれも希望者を対象に有料受験という形態で実施した。合計受験者数は、のべ504名(学部生466名、大学院生38名)で、うち80名(学部生78名、大学院生2名)は同一学生が複数回受験した。なお、複数回受験者の識別と非正規学生の扱いは、上記TOEIC IPの場合と同じである。

(3) TOEIC Speaking/Writing IPテスト

TOEIC Speaking/Writing IPテストは2011年11月と2012年12月の計2回実施し、合計受験者数は、32名(学部生28名、大学院生4名)で、複数回受験者はいなかった。

(4) スコアの集計と分析方法

収集したスコアは、言語教育センター専任教員である土肥と、当時本学の人文社会科学

研究科大学院生であった張が共同で学内において各種分析を行った。個人情報には十分注意し、まず氏名等の情報を削除して受験番号を使用して分析作業を行い、論文には受験番号等を含む個人が特定できる情報を一切掲載しなかった。なお、テスト実施の際にはスコアを統計目的で使用することを説明して了解を得ている。

より詳細な分析方法は、後述の4～7の「結果」欄で述べるが、本論文における数値は丸めの誤差を含む。また、表の見やすさを重視し、原則的に平均と標準偏差の値は小数点以下を四捨五入し、整数で示した。

4. TOEIC IP分析結果

以下では、土肥（2006）および土肥・柳瀬（2009）との内容の重複をできる限り避けるため、新たな知見が得られた結果を中心に掲載するが、一部のデータについては土肥・柳瀬（2009）（以下、「前回調査」と呼ぶ）との比較も行った。

(1) TOEIC IP学部・研究科別集計

TOEIC IP受験者21,280名の学部・研究科別のTotal, Listening, Readingの集計結果を表1に示した。ただし、学部名と研究科名は記号で示し、Totalの平均点が高い順に並べた。「最大値」「最小値」の数値の比較から、個人別のばらつきが極めて大きく、全学の最大値は理論的最大値（Total 990, Listening 495, Reading 495）と一致して満点を取った学生がいた半面で、理論的最小値（Total 10, Listening 5, Reading 5）と近い学生もいることがわかる。なお、各学部と研究科のListeningとReadingの最大値（または最小値）を足してもTotalの最大値（または最小値）と一致しない場合があるのは、ListeningとReadingの最大値（または最小値）を取った学生が異なることに起因する。このような極端な値も含めた上で平均を算出してみると、前回調査では2003～2007年度の5年分6,689名のデータでTotal, Listening, Readingの平均がそれぞれ506, 270, 235（丸めの誤差を含む：以下同様）であったのに対し、その後5年分を加えた今回の10年分21,280名の平均はそれぞれ505, 270, 235となり、非常に近い点数となった。2007年度以降は年1回1年次学生が受験義務を課されたことにより、意欲と英語力の低い学生が相当数含まれ、平均点が大きく下がることを予想していたが、年々1年次学生の平均点が上昇してきていることによってその影響は緩和され、むしろ意欲と英語力の高い学生の自主的な2回目以降の受験が増えているのかもしれない。

表1 TOEIC IP学部・研究科別集計

学部・研究科		受験者数	Total				Listening				Reading			
			平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値
学部	A学部	1,108	672	121	985	195	337	72	495	5	335	63	490	30
	B学部	1,073	604	110	975	210	309	64	495	75	295	60	495	75
	C学部	1,691	539	121	980	145	288	64	495	65	252	68	485	70
	D学部	3,528	531	112	990	15	277	59	495	5	254	65	495	5
	E学部	1,756	492	122	935	125	259	63	495	20	233	70	465	5
	F学部	5,491	485	122	990	75	262	65	495	5	224	68	495	5
	G学部	676	476	131	980	165	267	70	495	55	209	72	490	50
	H学部	1,814	459	119	910	75	253	62	495	45	206	67	430	10
	I学部	3,302	431	126	965	20	240	66	495	15	191	69	470	5
	全学部	20,439	503	134	990	15	269	68	495	5	235	76	495	5
研究科	J研究科	1	640	—	640	640	315	—	315	315	325	—	325	325
	K研究科	52	606	96	765	335	305	51	430	200	301	58	405	135
	L研究科	54	586	106	820	390	297	63	460	135	288	58	420	175
	M研究科	53	568	109	850	300	286	62	495	135	282	57	405	165
	N研究科	165	564	126	895	250	299	67	495	145	266	70	410	95
	O研究科	9	547	96	690	400	288	47	355	195	258	64	335	130
	P研究科	1	545	—	545	545	270	—	270	270	275	—	275	275
	Q研究科	119	531	133	860	235	291	68	455	135	239	74	415	80
	R研究科	30	519	170	950	235	283	84	465	150	236	92	485	70
	不明*	357	531	128	945	180	284	68	490	120	247	72	460	50
全研究科	841	548	128	950	180	290	67	495	120	258	73	485	50	
全学部・研究科	21,280	505	134	990	15	270	68	495	5	235	76	495	5	

*2003～2008年度受験の大学院生は研究科名を識別できないため、「不明」とした。

学部別の比較については、前回調査ではもっともTotalの平均が高い学部のTotal, Listening, Readingがそれぞれ644, 324, 320で、もっとも低い学部がそれぞれ440, 245, 195であったが、表1と比較すると格差が広がったことが判明した。この原因についても、意欲の高い学生と低い学生の差が広がっているものと考えられる。大学院生全体の比較については、前回調査のTotal, Listening, Readingの平均点が528, 282, 245であったのに対し、今回は上昇している。高得点で単位認定される場合がある学部生と異なり、大学院生は受験しても直接的なメリットはないが、英語力の向上を真剣に考える学生の受験が増えてきたのかもしれない。

近年の全国平均ではListeningがReadingより50点以上高いのに対し、本学の学部と研究科の平均ではそれぞれ34, 32点高いだけであることが判明した。前回調査の際にも同様の指摘をしたが、とくに平均点の高い学部（典型的な例はA学部）と研究科でその傾向が顕著である。1979年のTOEIC開始当時、（音声が軽視される英語教育を受けてきた）日本人のListeningとReadingの平均がそれぞれ約250点であったことを考慮すると、今ではListeningの点数のほうが50点以上高く出るほうが自然なことであり、本学の学生が入試対策のために文字英語にかなり偏った学習をしてきたことが影響しているのであろう。入学後の英語教育では入試対策ほどの偏りはないが、学生によっては音声よりも論文等の読み書きを重視する場合も考えられる。それが必ずしも悪いことであるとは言い切れないが、英

語をコミュニケーションの手段として考えた場合に、バランスの良い英語力が求められるのも当然のことである。

(2) TOEIC IP複数回受験者の上昇の傾向（学部別および全研究科）

前回調査でも複数回受験者の上昇の傾向について、「受験回数が多いほど上昇する傾向」と「初回点数が低いほど上昇する傾向」を示し、今回調査でも同様の傾向が再確認できたが、今回は詳細なデータの提示は省略する。前回調査では学部別の上昇量を示さなかったため、今回は表2に示した。なお、「複数回受験者の上昇量」の定義は前回と同じで、ひとりで2回以上（最高で15回）受験した学生の初回と最終回の得点差を示している。

表2の「学部・研究科」の列は表1と同一学部を同記号で表したが、大学院については複数回受験者が全体で90名と少なく、研究科別に分けても個人のばらつきの影響が大きいので、「全研究科」のみを示した。「間隔平均」は各受験者の初回から最終回までの日数の平均値を算出した。右6列にはTotal, Listening, Readingの初回平均と最終回平均を示した。それぞれの「最大値」「最小値」の列を見ると、各行において大きく得点を伸ばした学生と大きく得点を下げた学生の差が大きいことがわかる。これは前回調査の902名の複数回受験者全員の上昇量のグラフを示したときにも明らかになったことであるが、今回調査でも、入学時の英語力に大きな個人差があるだけでなく、入学後の上昇量で見ても大きな個人差があることがわかる。しかし、学部別や全学の平均を取って見るとどの学部も上昇を示し、全体としては学習効果が表れていることも判明した。

表2 TOEIC IP複数回受験者の学部別上昇量

学部・研究科	受験者数	回数平均	間隔平均	Total上昇				Listening上昇				Reading上昇				Total		Listening		Reading	
				平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値	初回	最終	初回	最終	初回	最終
A 学部	157	2.2	207	38	109	375	-435	9	74	195	-275	29	60	235	-310	609	647	309	318	300	329
B 学部	189	2.7	697	49	95	385	-305	17	56	170	-240	32	60	215	-175	563	612	290	307	273	305
C 学部	210	2.6	627	72	108	445	-215	35	62	265	-155	37	62	260	-145	549	621	293	328	257	293
D 学部	403	2.7	550	49	91	525	-285	22	52	265	-280	27	56	260	-135	537	586	280	302	258	284
E 学部	156	2.7	594	79	114	730	-280	33	65	300	-215	46	68	430	-215	486	565	261	294	225	271
F 学部	496	2.5	608	66	98	410	-210	30	56	235	-185	35	61	255	-160	495	561	267	297	229	264
G 学部	52	2.3	311	55	108	380	-310	24	53	175	-140	31	68	205	-170	493	548	282	306	211	242
H 学部	226	2.6	714	82	108	635	-160	35	61	295	-140	47	62	340	-80	457	538	255	290	202	248
I 学部	193	2.5	567	75	99	500	-130	40	53	270	-70	35	62	340	-100	494	569	270	310	224	259
全学部	2082	2.6	576	63	102	730	-435	28	59	300	-280	35	61	430	-310	519	581	276	304	242	277
全研究科	90	2.3	240	43	75	290	-125	21	44	145	-90	22	51	200	-85	559	602	293	314	266	288
全学部・研究科	2172	2.6	562	62	101	730	-435	28	58	300	-280	34	61	430	-310	520	582	277	304	243	278

表2で学部間の比較をすると、興味深いことが見えてくる。表1の説明で、全国的にはListeningがReadingより50点以上高いはずが、A学部をはじめとする得点の高い学部ほどListeningとReadingの点数の差が小さい傾向を示していることを述べたが、表2の上昇量の方でも英語力が高い学部ほど、在学中のReadingの上昇量がListeningの上昇量よりも高い傾向がある。もともと偏りのあった英語力に、より大きな偏りが加わったというこ

とが懸念される。

表2から全体の平均では、562日の間隔でTotal, Listening, Readingの点数がそれぞれ62, 28, 34伸びたことがわかる。これを「一日当たりの上昇量」に換算したものが表3である。後述するように、この数値を用いて他のテストとの比較を行う。

表3 TOEIC IP Section別の一日当たりの上昇量 (全学部・研究科)

Section	1日当たりの上昇量
Total	0.11
Listening	0.05
Reading	0.06

また、表2を見ると、英語力をもっとも高いA学部のTotal上昇量をもっとも低い(38点)ということになるが、Totalの学部別上昇量を図示すると図1のようになり、受験間隔が短い割にはよく上昇しているとも言える。他学部についても多少のばらつきがあるとは言え、学部別の平均を比較するとほぼ同様の上昇傾向を示している。

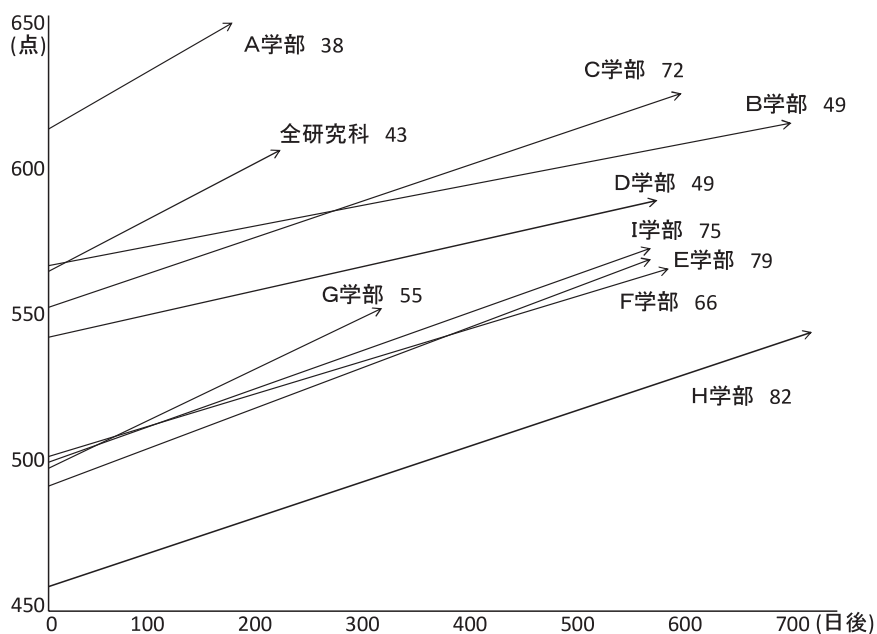


図1 学部別TOEIC IP Totalスコアの推移 (学部・研究科の数値は上昇量)

(3) TOEIC IP複数回受験者のTotal上昇の傾向 (1年次初回からの推移)

前回調査では1年次から4年次までのどの組み合わせでも、平均すればTOEIC IPのスコアが下降せずに上昇していることを示したが、今回は少し観点を換え、1年次全学生が前期に義務として受験したときのスコアを基点として、その後どれだけTotalのスコアが

伸びているかを示すこととした。

1年次全学生が前期に受験した日時と平均点は、2007年8月1日（429点）、2008年6月7日（458点）、2009年6月6日（499点）、2010年5月29日（494点）、2011年6月4日（495点）、2012年6月2日（501点）であった。このうち初回の2007年はとくに平均点が低かったが、(1) 前期授業終了翌日で、(2) 猛暑であった、という2つの明確な理由があり、今回の「基点」としての分析からは除外することとし、それ以外の2008～2012年の6月の受験回を基点として採用することとした（2010年は5月29日であったが、他の年度の6月と大きな差はないと判断し、ここでは便宜上6月実施と呼ぶ）。つまり、2008～2012年6月に1年次学生がそれぞれTOEIC IPを受験し、そのうちの一部学生が他の時期にも自主的に受験したということになる。たとえば、1年次6月に受験して1年次7月にも受験した学生が19名いたが、その19名の中の6月と7月の平均点はそれぞれ554点と588点であり、1か月で34点上昇したことになる。同様に、1年次6月と1年次10月に受験した学生が71名いて、その71名の中の6月と10月の平均点はそれぞれ565点と606点であり、4か月で41点上昇したことになる。このような考え方で、いずれも1年次6月を基点として、その後の受験時に同一学生同士の比較でどれだけ上昇したかを図示したのが図2である。5年目以降の受験者数は少ないので省略したが、4年目までの各回についてはもっとも人数が少なかったのが4年次2月との組み合わせで14名であった。

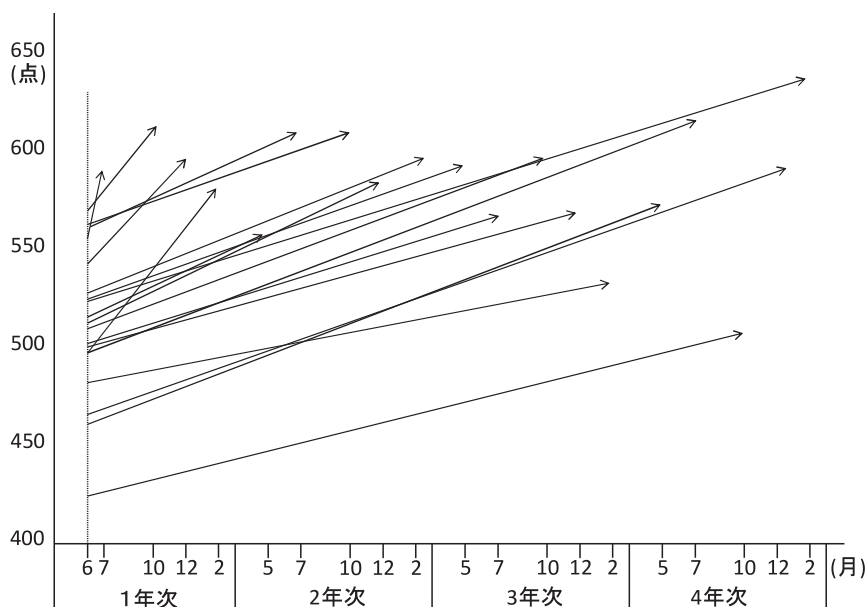


図2 TOEIC IP Total の実施回別推移

図2を見ると、1年次の間は矢印の傾きが急で、2年次以降は緩やかになっていることが明らかで、学生が1年次に卒業単位をそろえるために英語の授業を集中して取り、2年

次以降の英語学習量が減少していることが大きく影響していると言える。

5. TOEFL ITP分析結果

これまで本学における TOEFL ITP の結果を包括的に分析した結果を公表してこなかったが、以下では過去6年間（2007年12月～2012年12月）にわたって、年1～2回計11回実施されたTOEFL ITPの結果を記す。

(1) TOEFL ITP学部別・全研究科集計

全受験者504名の学部別のTotal, Section 1（聴解）, Section 2（構文／文法）, Section 3（読解）の集計結果を表4に示した。ただし、学部記号は表1のTOEIC IPと同じ記号を用い、学部の並び順も表1に合わせてTOEIC IPのTotalの平均点が高い順にした。大学院生についてはN研究科の受験者が最大の13名で、他の研究科はそれぞれ10名未満となり、研究科別に分ける意義が小さいと判断し、全研究科のデータのみを示した。

表4 TOEFL ITP学部別・全研究科集計

学部・研究科	受験者数	Total				Section 1（聴解）				Section 2（構文／文法）				Section 3（読解）			
		平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値
A 学部	4	524	22	550	497	52	3	55	47	51	6	59	43	55	1	56	53
B 学部	3	517	14	537	507	49	2	51	47	52	3	56	50	53	2	55	50
C 学部	56	488	33	563	410	48	4	58	41	48	4	59	39	50	4	58	37
D 学部	58	491	35	610	390	48	4	64	40	49	4	58	38	50	4	61	37
E 学部	44	487	36	573	420	46	5	60	31	49	5	60	39	51	4	62	43
F 学部	159	477	37	593	373	47	4	58	34	47	5	61	31	49	5	61	31
G 学部	5	452	51	547	397	46	6	57	39	44	4	53	41	45	6	54	37
H 学部	27	463	35	573	397	46	4	59	38	46	4	55	38	48	4	58	42
I 学部	110	493	40	590	373	49	4	57	36	49	5	63	35	50	5	61	37
全学部	466	484	38	610	373	48	4	64	31	48	5	63	31	50	5	62	31
全研究科	38	488	43	567	320	49	4	58	33	48	6	59	32	50	6	59	31
全学部・研究科	504	485	39	610	320	48	4	64	31	48	5	63	31	50	5	62	31

TOEFL ITPは旧形式の TOEFL PBT（Paper-Based Test）と同形式、同難易度のテストであるが、前述のTOEICのように30年以上前の日本人を基準にしてスコアの絶対基準が定められたものではなく、世界基準で定められたものである。各Sectionの最小値は31、最高値は67～68で、Totalの最小と最大はそれぞれ310、677で、2012年の日本語を母語話者とする受験者の平均はSection 1, Section 2, Section 3, Totalがそれぞれ54, 54, 53, 536である。このような日本人の平均を参照しながら表4のSection 1 と Section 3 に注目すると、千葉大生の聴解力が相対的に低い傾向と、英語力が高い学部の方がその傾向が顕著である傾向が見えてくる。学内におけるTOEIC IPとTOEFL ITPの受験者層は異なり、TOEIC IPは全学生が受験し、TOEFL ITPは留学志向の比較的高い自主的な受験者がほとんどであって、まったく同等の比較をすることはできないが、それでもTOEICとTOEFLという別の観点で見ても聴解力が弱いという同様の結果が示されたことは興味深い。今後のカリキュラムの開発や学習目標の設定にも、この結果を生かす必要がある。

(2) TOEFL ITP複数回受験者の上昇の傾向

前述のTOEIC IPと同様にTOEFL ITPの上昇傾向についても、学部別・研究科別の分析を行った。しかし、複数回受験した学生が9学部中3学部で10名未満となることと、大学院生については全研究科を合わせても該当者が2名しかいないことも判明した。本論文では誤解を避けるため、学部別・研究科別の結果や学年別の結果を示さず、全学の該当者80名の平均のみを表5に示した。

表5 TOEFL ITP複数回受験者の上昇量

学部・研究科	受験者数	回数平均	間隔平均	Total上昇				Section 1上昇				Section 2上昇				Section 3上昇			
				平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値
全学部・研究科	80	2.4	381	18	28	123	-44	2	4	18	-13	2	4	13	-7	1	4	11	-8

TOEIC IPの複数回受験者が2,172名いたことと比べると、TOEFL ITPの複数回受験者は80名と少ないが、それでもある程度信頼に足る数であると考ええる。表5の通り、381日の間隔でTotal, Section 1, Section 2, Section 3 の上昇量がそれぞれ18, 2, 2, 1であり、これも個人差（最大値と最小値の差や標準偏差）は大きいですが、全体として英語力を伸ばしていることが確認できた。各Section の1～2という上昇量が感覚的に小さいように見えるが、TOEFL ITPのTotalのスコアは、各Sectionの平均の10倍であることを考えると、実際には「10倍の価値」があるとの言い方も可能である。この上昇量がTOEICとの換算でどの程度の意味を持つのかについては、後述する。

6. TOEIC Speaking/Writing IPテスト分析結果

通常のTOEIC公開テストや同形式のTOEIC IPが直接測定しているのが、ListeningとReadingの受信技能だけであるとの批判を受けて2006年に開始されたのがTOEIC Speaking/Writingテストであり、本学ではIP(団体特別受験制度)を2011年11月と2012年12月に計2回試験的に実施した。

表6 TOEIC Speaking/Writing IPテスト集計 (千葉大生)

Score / Level	受験者数	平均	標準偏差	最大値	最小値
Speaking Score	32	113	18	150	70
Writing Score	32	137	22	170	80
Speaking Proficiency Level	32	5	1	6	3
Pronunciation Level	32	2	0	3	1
Intonation and Stress Level	32	2	0	3	1
Writing Proficiency Level	32	7	1	8	4

表6のうち、各受験者のScore はそれぞれ0～200点の10点刻みで示され、Speaking

Proficiency Levelは1～8の8段階で、Pronunciation LevelとIntonation and Stress Levelは1～3の3段階で、Writing Proficiency Levelは1～9の9段階で示される。各段階の詳細な位置付けは国際ビジネスコミュニケーション協会（2013b）を参照されたいが、公表されているSpeaking ScoreとWriting Scoreの全国データを国際ビジネスコミュニケーション協会（2013c）から抜粋したのが表7である。

表7 TOEIC Speaking/Writing公開テストおよびIPテスト全国平均（学生・社会人）

TOEIC Speaking/ Writing テスト	学 生				社 会 人			
	公開テスト		IPテスト		公開テスト		IPテスト	
	受験者数	平 均	受験者数	平 均	受験者数	平 均	受験者数	平 均
Speaking Score	1,687	123	2,475	104	3,207	125	2,475	117
Writing Score	1,687	136	1,832	117	3,207	139	1,973	134

(TOEICプログラムDATA & ANALYSIS 2012より抜粋)

校内においても学外においても、Speaking/Writingテストの人数や受験者層はTOEICとは異なるが、表6と表7を比較すると、発信技能においても千葉大生は音声英語（Speaking）が文字英語（Writing）よりも弱い傾向が伺える。

7. 異なるテスト間のスコア比較の試みと結果

本論文で取り上げた以外のものも含め、英語テストには複数の種類があり、それぞれに長所と短所があるため、いずれかのテスト一種類だけで十分ということにはならない。しかし、英語力を示す尺度として複数のものが並列している場合に、それらの比較や換算ができないと教育や研究の面で不便なことがある。たとえば千葉大学では、TOEIC（公開テストとIPテスト）、TOEFL（iBTとITP）、IELTSの基準を用いて高得点者に単位認定をしたり、中級英語、上級英語の履修資格を定めたりしている。その基準を示す資料（千葉大学言語教育センター, 2013）の一部を抜粋したのが表8である。しかし、この換算基準は他の教育・研究機関の資料や教員の主観も織り交ぜて数年前に決定したものであり、必ずしも厳密な科学的根拠に基づいて決められたものではなかった。Educational Testing Service はかつて $TOEFL = TOEIC \times 0.348 + 296$ という換算式を公表していたが、現在では撤回されている。その後 Educational Testing Service (2003) はTOEIC で50点刻みの範囲が TOEFLスコアの何点の範囲に相当するかという目安の表を示したが、受験者ひとりひとりの得点を換算式として示すような性質のものではない。上記の本学の目的に適した形で、公開された換算式が存在しないのが現状であり、本研究では大量に収集したスコアを使用して換算式を作成することを目指した。

表8 異なるテスト間の換算の例

単位認定基準(平成22年度以降入学の学部生対象)(TOEFL iBT/TOEIC/IELTS の基準も併記します):					
認定科目	TOEFL ITP, TOEFL PBT	TOEFL iBT	TOEIC (IP)	IELTS	
検定英語 I (2単位)	530~549点	71~78点	680~729点	該当なし	
検定英語 I・II (計4単位)	550~677点	79~120点	730~990点	6	
授業受講資格 (TOEFL iBT および TOEIC の基準も併記します):					
授業科目	TOEFL ITP, TOEFL PBT	TOEFL iBT	TOEIC (IP)	IELTS	
中級英語 I (履修後に2単位)	487点以上	57点以上	550点以上	5以上	
中級英語 II (履修後に2単位)	523点以上	69点以上	650点以上	5.5以上	
上級英語 (履修後に2単位)	550点以上	79点以上	730点以上	6以上	

(「平成25年度 TOEFL ITP学内実施のお知らせ」より抜粋)

(1) TOEIC IP Total と TOEFL ITP Totalの比較

TOEIC IPの受験者がのべ21,280名で、TOEFL ITPの受験者がのべ504名であったことを述べたが、単純に言えば両方を受験した学生のみを抽出し、片方を横軸に、もう片方を縦軸にして散布図を描き、回帰直線を引けば換算式を作成することが可能である。しかし、同一学生であってもTOEIC IPとTOEFL ITPの受験の間にたとえば3年もの間隔があるとすれば、その間に英語力が大きく変動することが考えられ、比較の妥当性が疑問視される。逆に、両テストの受験間隔をたとえば1週間と極端に狭く限定してしまうと、そんな短期間に両方のテストを受験した学生は極めて少人数であり、比較の信頼性が疑問視される。このような考えに立脚し、受験間隔と該当者数のバランスを考慮してデータの抽出作業を何度か試行した結果、受験間隔を2か月以内に設定した場合、同一学生がTOEIC IPとTOEFL ITPの両方を受験した一対一の対応があるデータセットが138件抽出できることが判明し、この138件を用いて分析を行うことにした。なお、ここで言う「2か月以内」の間隔とは、TOEIC IPと TOEFL ITPのいずれが先でもよく、また受験日の年と月だけを比較し、月内の日付までは考慮しないという基準を用いた。たとえばある学生が2011年12月17日にTOEFL ITPを受験した場合、同一学生が2011年10月1日から2012年2月29日までにTOEIC IPを1回受験していれば、一対のデータセットが見つかったということになる。つまり、受験間隔は「正味2か月以内」ではなく、「2か月強」の可能性も含まれるということである。また、上記の例で同一学生が2011年10月1日から2012年2月29日までにTOEIC IPを2回受験していれば、二対のデータセットが見つかったということになる。

このような基準で抽出した138対のデータセットを用い、まず相関(暫定値)を算出したところ、 $r = .747$ という高い値が得られた。この分析方法に対しては、2か月以内とは言え、その間に若干の英語力の変動があるのではないかとの指摘も予想される。そのよう

な指摘に対しては、TOEIC IPが先の場合も後の場合もあり、両テストの受験の間に英語力が上がる場合も下がる場合もあり、138件ものデータセットがあれば英語力の変化の影響は相殺できるとの反論も可能ではある。しかし本研究では、少しでも分析の精度を上げることを目指し、表3に示したTOEIC IP Totalの一日当たりの上昇量 (0.11) を使って補正をすることにした。たとえば、TOEIC IP受験の30日後にTOEFL ITPを受験した場合、TOEFL ITPの受験当日にはTOEIC IPのTotalスコアが3.3点上昇しているはず (+0.11×30) という前提である。別の例を挙げれば、TOEFL ITP受験の45日後にTOEIC IPを受験した場合、TOEFL ITPの受験当日にはTOEIC IPが4.95点低かったはず (-0.11×45) という前提である。このような補正を行ってから再度138対のデータセットの相関を算出したところ、 $r = .750$ となり、上記の $r = .747$ よりわずかではあるが改善している（少なくとも改悪していない）ことが判明した。138対のデータセットについて、補正後のTOEIC IP Totalを横軸、TOEFL ITP Totalを縦軸にして描いた散布図を図3に示した。この補正後のデータに基づいて単回帰分析による換算式と前述の補正前後の相関を表9に示した。

TOEFL-Total / TOEIC-補正Total

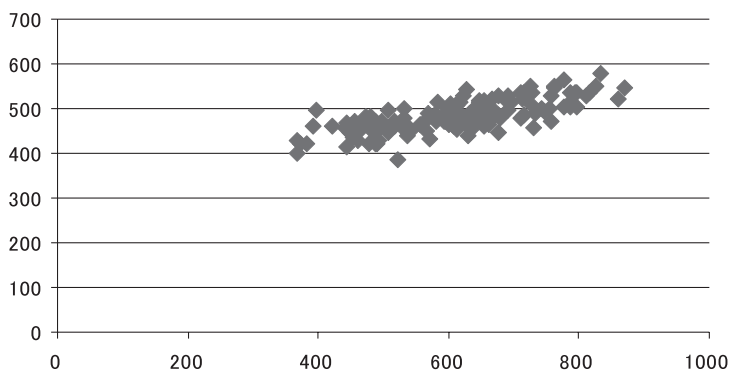


図3 TOEIC IP Total補正スコアとTOEFL ITP Totalスコアの散布図

表9 TOEIC IP Total と TOEFL ITP Total の関連

データセット数	138
相関 (補正なし)	$r = .747$
相関 (補正あり)	$r = .750$
換 算 式	$[\text{TOEFL ITP Total}] = [\text{TOEIC IP Total}] \times 0.229 + 344$

表9の換算式を用いて表8の換算表の妥当性を検証してみると、たとえば表8ではTOEIC IP 680点は TOEFL ITP 530点に相当することになるが、表9では TOEIC IP 680点は TOEFL ITP で約500点に相当することになる。以前から複数の学生、教員等から

TOEFL の基準がTOEICの基準より難しいとの指摘を受けていたこととも整合する。本研究の結果に基づいて、いずれ換算表を修正する必要があるが、これまで数年間使用してきた基準を急に変更することは教育現場の混乱にもつながりかねないため、次回何らかの制度変更をする際に同時に換算表も作り変えるのが妥当と考える。

(2) TOEIC IPとTOEFL ITPの下位Sectionの比較

上記の TOEIC IP Total と TOEFL ITP Total の関連性の比較をしたのと同じ考え方にに基づき、TOEIC IP Listening Section と TOEFL ITP Section 1（聴解）の比較、および TOEIC IP Reading Section と TOEFL ITP Section 3（読解）の比較を行った。ただし、スコアの補正については、表3に基づいて TOEIC IP Listening と Readingがそれぞれ一日当たり0.05点、0.06点上昇するという前提を用いた。結果は図4、図5、表10、表11に示した。

TOEFL-S1 / TOEIC-補正L

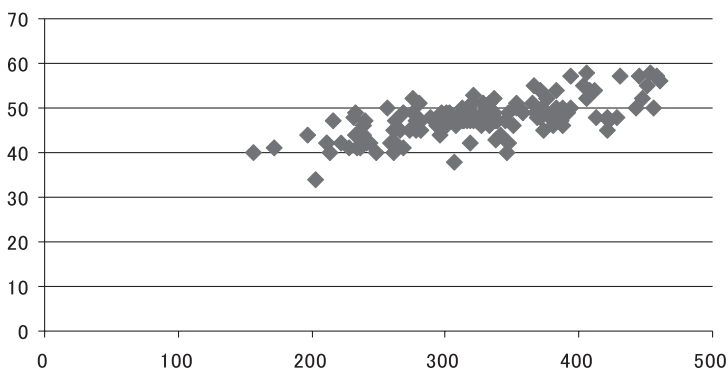


図4 TOEIC IP Listening補正スコアとTOEFL ITP Section 1（聴解）スコアの散布図

TOEFL-S3 / TOEIC-補正R

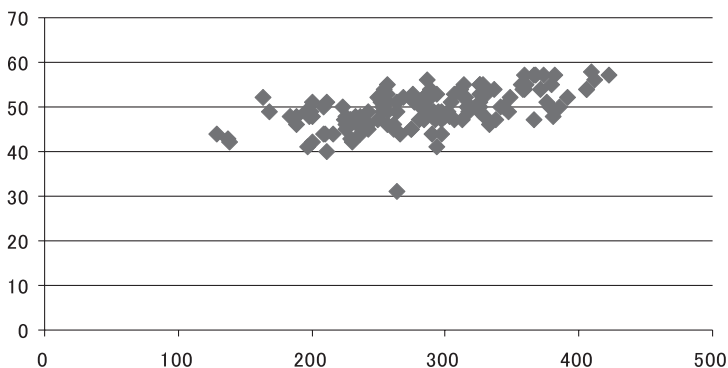


図5 TOEIC IP Reading補正スコアとTOEFL ITP Section 3（読解）スコアの散布図

表10 TOEIC IP Listening と TOEFL ITP Section 1 (聴解) の関連

データセット数	138
相関(補正なし)	$r = .687$
相関(補正あり)	$r = .688$
換 算 式	$[\text{TOEFL ITP Section 1}] = [\text{TOEIC IP Listening}] \times 0.044 + 33$

表11 TOEIC IP Reading と TOEFL ITP Section 3 (読解) の関連

データセット数	138
相関(補正なし)	$r = .579$
相関(補正あり)	$r = .584$
換 算 式	$[\text{TOEFL ITP Section 3}] = [\text{TOEIC IP Reading}] \times 0.040 + 38$

図4と図5を見ると、図3よりもデータの収束性が低いように見え、その結果が表10と表11の相関にも表れているようである。現状では大学としてListeningやReadingの能力についてTOEICとTOEFL間で換算する必要性は高くないが、各学生のテスト結果を解釈する際にはこのようなデータも参考になるはずである。

(3) TOEIC IPとTOEIC Speaking/Writing IPテストの比較 (参考)

前述の通り、TOEIC IPにはTotal, Listening, Reading の3種のスコアがあり、Speaking/Writing IPにはSpeakingとWritingの2種のスコアがあり、このうちのいくつかの組み合わせについて、すべてを受験した28名のデータを使用して関連性を調査して図6に散布図を、表12に相関と換算式を示した。ただし、受験者数が少ないこともあり、あくまでも参考値として、前述の2か月以内の受験間隔という制限は設けず、表3の一日あたりのTOEIC IPの上昇量に基づく補正を行った。

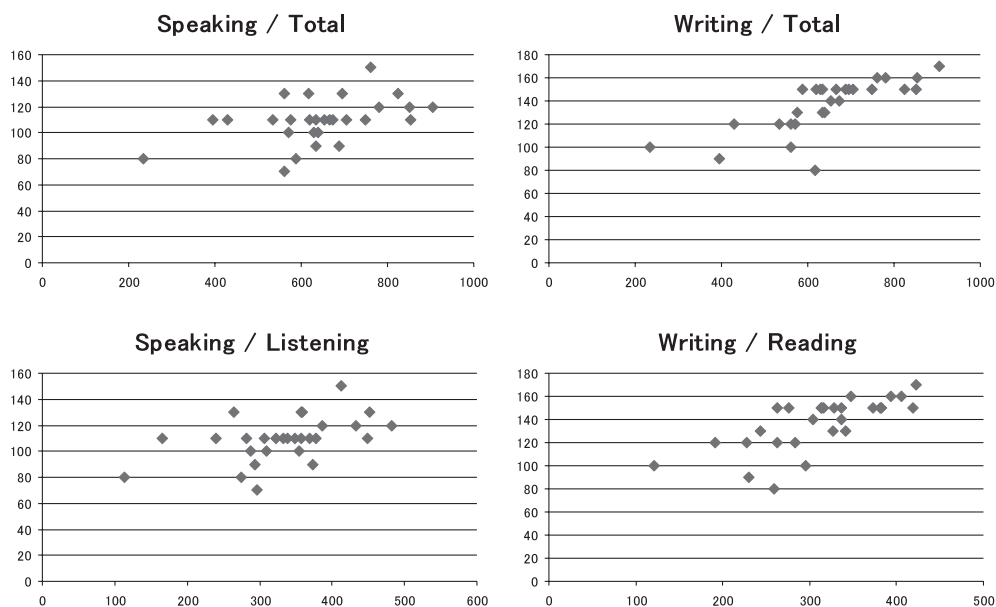


図6 TOEIC IP とTOEIC Speaking/Writing IPテストのSection別の散布図（4種）

表12 TOEIC IP とTOEIC Speaking/Writing IPテストのSection別の相関（4種）

Section組み合わせ	相関（補正なし）	相関（補正あり）	換算式（参考）
Speaking/Total	r = .416	r = .459	[Speaking]=[TOEIC Total]×0.055+74
Writing/Total	r = .746	r = .758	[Writing]=[TOEIC Total]×0.120+59
Speaking/Listening	r = .447	r = .488	[Speaking]=[Listening]×0.103+75
Writing/Reading	r = .728	r = .779	[Writing]=[Reading]×0.681+83

国際ビジネスコミュニケーション協会（2007）には、日本と韓国の2,064名の両テスト受験結果に基づいて、TOEICスコアからSpeaking/Writingスコアを予測するための「比較表」が用意されている。しかし、両テストを受験した28名の千葉大生は、この比較表の予測よりも低いSpeaking/Writingスコアが出る傾向が出るようになった。2007年作成の換算表に問題があるのかどうかは不明であるが、今回の千葉大のデータが28名と少なく信頼性が不十分であることは事実で、また、千葉大生の発信能力が決して高くはないことは日頃の経験からも感じられることである。

8. まとめ

本研究では、TOEIC IPについて前回調査よりも大規模かつ長期的なデータに基づいて、過去の結果を再確認するだけでなく、新たな分析方法によって千葉大学における英語学習の成果と問題点を示すことができた。またTOEFL ITPとTOEIC Speaking/Writing IP

Testsについては、これまで学外においても十分な研究がなかったが、本研究の結果によりスコアの相関と換算についての知見が得られた。本研究の個々のデータについての考察はすでに述べたが、全体として以下のようなことが言えると考ええる。

まず、千葉大生の英語力には大きなばらつきがあるが、平均して見ると十分な到達レベルに達していないということである。このような意図もあって、今回はそれぞれの上昇データについて統計的検定で有意差の有無を調査しなかった。TOEIC IPの平均505点（表1）は「どんな状況でも適切なコミュニケーションができる素地を備えている」と評価される730点のレベルには遠く及ばない。Listening 270点、Reading 235点もCEFR（ヨーロッパ言語共通参照枠）ではA2レベルのBasic Userに相当する。比較的英語力の高い学生が受けたと考えられるTOEFL ITPの平均485点（表4）も留学に必要なレベルではない。Speaking 113点（表6）はTOEICの評価では「日常のコミュニケーションレベル」で、CEFRではA2レベル（Basic User）である。もう一方のWriting 137点（表6）も「日常のコミュニケーションレベル」と「海外出張レベル」の中間ほどであり、CEFRではB1（Independent User）の下のほうに位置付けられる。また複数のテストデータを総合すると、千葉大生は平均的な日本人と比べて相対的に、ReadingよりListeningが苦手で、受信（ListeningとReading）より発信（とくにSpeaking）が苦手である傾向が示唆された。

次に、平均では英語力が不十分なレベルであるとしても、学部別のTOEIC IPの上昇量（図1）や学年別のTOEIC IPの上昇量（図2）では統計的検定を行わずともすべてのケースで上昇傾向が見られ、学生たちの努力の成果が表れていると言える。前回調査でも述べたように、複数回受験する学生は、比較的意欲が高いと考えられ、全学生に強制的に複数回受験させればもっと悪い結果が予想できる。しかし、やる気を出せば英語力が伸びるという当然の結果を示せたことは意義がある。TOEIC IPが562日で62点上昇した（表2）のに対し、TOEFL ITPは381日で18点上昇した（表5）。この18点が大きいかどうかについては、表9の換算式を利用すれば、TOEICでは79点程度（ $=18/0.229$ ）の上昇に相当することになり、見かけ上の数字の小ささに反して、相当な上昇量であることもわかる。全学生がグローバル社会の即戦力とは言えなくとも、少なくとも潜在力を伸ばすことに成功していると考ええる。

本研究は学生の英語力の実態を示すのみならず、これまで十分な先行研究がなかったTOEICとTOEFL（さらにはSpeaking/Writing）の換算式を作成する研究も行った。異なるテストの受験間隔にもある程度の英語力の変動があることを予測して得点を補正するという手法を試みたが、表9～12のいずれでも補正後は相関が高くなっていることが判明し、補正することの妥当性を示せた。換算式作成にあたり、ある程度の人数のデータを集めることができたが、すべて千葉大生のデータであるので、学外でも汎用性のある換算式であるかどうかは検証の余地がある。

テストだけ実施していても英語力が伸びないのは当然のことであり、今回のデータが示したように日頃の努力が英語力の上昇に結び付くのは言うまでもない。今後もテストのデータを収集するだけでなく、日頃の指導方法・学習方法との関連性を調査することによ

り、指導と評価の両輪を充実させる必要がある。

謝辞

本論文の執筆にあたり、TOEIC IPと TOEFL ITPの実施のためにご指導とご協力をいただいた千葉大学の歴代の学長、言語教育センター長、普遍教育センター長、英語教員の皆様、関係各部署の事務担当職員の皆様、学生の皆様、千葉大学生協同組合、国際ビジネスコミュニケーション協会、国際教育交換協議会（CIEE）日本代表部、その他の千葉大学内外の多くの方々に感謝の意を表します。

主な参考文献

- 千葉大学言語教育センター, 「平成25年度 TOEFL ITP学内実施のお知らせ」, <http://f.chiba-u.jp/english/documents/2013toeflitp.pdf>, 2013.
- 土肥充, 「TOEIC IPによる千葉大生の英語力の現状分析」, 『人文と教育』, 千葉大学国際教育開発センター, 第2号, pp.15-29, 2006.
- 土肥充, 柳瀬弘美, 「千葉大学におけるTOEIC IPスコアの包括的分析」, 『言語文化論叢』, 千葉大学言語教育センター, 第3号, pp.31-45, <http://f.chiba-u.jp/about/plc03/plc03-06.pdf>, 2009.
- Educational Testing Service, “TOEFL Institutional Testing Program (ITP) and TOEIC Institutional Program (IP): Two On-Site Testing Tools from ETS at a Glance”, 2003.
- Educational Testing Service, TOEFL, <http://www.ets.org/toefl/>, 2013a.
- Educational Testing Service, Test and Score Data Summary for TOEFL iBT Tests and TOEFL PBT Tests: January 2012-December 2012 Test Data, http://www.ets.org/s/toefl/pdf/94227_unlweb.pdf, 2013b.
- Educational Testing Service, TOEIC, <http://www.ets.org/toEIC/>, 2013c.
- Educational Testing Service, TOEIC Speaking/Writing Tests: About the Tests, http://www.ets.org/toEIC/speaking_writing/about/, 2013d.
- グローバル人材育成推進会議, 「グローバル人材育成戦略（グローバル人材育成推進会議審議まとめ）」, <http://www.kantei.go.jp/jp/singi/global/1206011matome.pdf>, 2012.
- 国際ビジネスコミュニケーション協会, 「TOEICテストとTOEICスピーキングテスト・ライティングテストのスコア比較表」, http://www.toEIC.or.jp/library/toEIC_data/sw/pdf/score_comparison_list.pdf, 2007.
- 国際ビジネスコミュニケーション協会, 「TOEICテスト受験者数の推移」, http://www.toEIC.or.jp/library/toEIC_data/toEIC/pdf/about/transition.pdf, 2013a.
- 国際ビジネスコミュニケーション協会, 「TOEICスピーキングテスト/ライティングテスト」, <http://www.toEIC.or.jp/sw.html>, 2013b.
- 国際ビジネスコミュニケーション協会, 「TOEICプログラムDATA & ANALYSIS

2012], 2013c.

Tannenbaum, Richard J. and E. Caroline Wylie, "Mapping English Language Proficiency Test Scores Onto the Common European Framework", *TOEFL Research Reports*, RR-80, 2005.

Tannenbaum, Richard J. and E. Caroline Wylie, "Linking English-Language Test Scores Onto the Common European Framework of Reference: An Application of Standard-Setting Methodology", *TOEFL iBT Research Report*, TOEFLiBT-06, 2008.