

# 千葉大学における TOEFL ITP のスコア分析

土肥 充

千葉大学国際教養学部

## An Analysis of TOEFL ITP Scores at Chiba University

Mitsuru Doi

### 要旨

本研究は、2016年度から千葉大学の新生を対象に実施することとなった TOEFL ITP について、2016年4月の1回分、2,359名への実施結果について分析したものである。得点の分布、学部別平均点の比較、学科別平均点の比較のいずれを見てもばらつきが大きいことが判明した。全学の平均点は461点で全国の大学生平均に近い数字であったが、CEFRのB2レベル以上に相当する543点以上が3.1%で、正規留学等が可能な実用レベルに達している学生は少ないと結論した。3つのSectionの比較をしてみると、全学的に同じ傾向を示すのではなく、学科別の傾向が異なることが判明した。また、TOEIC L&R IP を全学的に実施した場合と比較して、TOEFL ITP ではSection別の理論上の最低点を取る学生が多いことを問題点として指摘した。最後に、土肥・張（2014）が提案した TOEFL ITP と TOEIC L&R IP の換算式について、今回のデータと過去のデータも比較しながら多面的に検証し、妥当性と問題点について論じた。

キーワード

TOEFL ITP、TOEIC L&R IP、換算式、CEFR

## 1. はじめに

グローバル化や説明責任等を求める社会の動きの中で、外部テストによる客観的な英語力の測定が、留学や就職のみならず、大学内の単位認定、英語授業のプレースメント、入試等でも必要とされる機会が増えてきた。TOEFL、TOEIC、IELTS、英検等の各種英語テストには、それぞれ異なる目的や特徴があり、どのテストを選択するのもかも学生や大学の目的や方針によって当然異なる。千葉大学では、学生が各自の希望によって学外で受験する複数の英語テストに対応して単位認定やプレースメントの基準を設定してきたが、学内でも2003年度から全学を対象にTOEIC® L&R(従来の TOEIC は2016年8月からTOEIC Listening & Reading と改称されたため、既刊の論文名等は除き、原則的に本論文での記述は過去に実施した分も改称後の略称である TOEIC L&R を使用する)の団体特別受験制度であるTOEIC L&R IPを実施し、2007年度からはTOEFL®の団体向けテストであるTOEFL ITP を実施してきた(土肥・張2014)。また、2007年度からは新入生のTOEIC L&R IPの受験が義務化され、2016年度からはTOEIC L&R IPのかわりにTOEFL ITPを新入生が受験することが義務化され、さらに学部によって2年次以降にTOEIC L&R IPまたは TOEFL ITP の受験が義務化されている。

TOEFL ITPやTOEIC L&R IP の問題形式や問題例については、テスト開発団体である米国ETS (Educational Testing Service) のウェブサイトや多数の参考書・問題集等で解説されているため、本論文では詳細な説明を省略するが、TOEFL はもともと英語圏留学の資格試験が本来の目的であったのに対し、TOEICは一般の日本人やビジネスマンも主要なターゲットとして開発されたもので、学術を主目的とするものではなく、使用される語彙や内容の傾向も異なる。また、TOEIC L&R IP は

Listening Section	5～495点
Reading Section	5～495点
Total(両Sectionの合計)	10～990点

の2つのSectionからなるのに対し、TOEFL ITP は下記3つの Section に分かれている(国際教育交換協議会2016b)。

Section I	Listening Comprehension	31～ 68点
Section II	Structure and Written Expression	31～ 68点
Section III	Reading Comprehension	31～ 67点
Total	各Sectionの平均を10倍し四捨五入	310～677点

TOEFL ITP には TOEIC L&R IPと類似する技能を測定するSection I とIIIに加え、Section IIが存在することにより、得られる知見も異なるはずである。

筆者はこれまで新入生対象のTOEIC L&R IP や TOEFL ITPの結果を集計して掲示や会議資料として学生や各学部教員に提供してきたが、それと並行して筆者らは千葉大学生協で希望者が受験できる TOEIC L&R IP やTOEFL ITP の結果も含めて、より詳細な分

析を行ってその結果を公刊してきた。土肥（2006）は2003～2004年度の2年分のべ1,240名の TOEIC L&R IP 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 TOEIC L&R IP スコアについて包括的分析を行った。さらに土肥・張（2014）は2003～2012年度の10年分のべ21,280名の TOEIC L&R IP の結果と合わせて、2007～2012年度の6年分のべ504名の TOEFL ITP の結果、さらには 2011～2012年度の2年分32名の TOEIC Speaking & Writing IP の結果も含めて、経年調査と比較調査を行った。以上3点の論文では、ただ単にデータが徐々に蓄積されてきたというだけでなく、異なった観点での分析を工夫し、英語テストから得られる幅広い知見を示してきた。本研究では、TOEFL ITP の新入生の受験義務化が始まってから1回分のデータしかないが、2016年4月に実施した2,359名の受験結果について分析する。筆者らの過去のTOEIC L&R IP の研究では千葉大学生協で自主的に受験したデータも含め複数年度にわたって同一学生の追跡調査をして、得点の上昇量を測定することを主要なテーマとしたが、今回の新入生の TOEFL ITP 受験についての研究では、論文執筆時点（2016年11月）で同一学生が別の機会に生協で自主的に複数回受験したケースがほとんど見当たらないため、今回は上昇量についての研究はせず、生協での自主受験のデータを含めないこととした。

## 2. 研究の目的

本研究の目的は、2016年4月に全学の新入生に対して実施した TOEFL ITP の結果を多面的に分析することである。

## 3. 研究の方法

本研究は、千葉大学で2016年4月9日に新入生2,359名を対象に実施した TOEFL ITP の個人別スコアを集計し、分析した。千葉大学には10の学部があり、そのうち国際教養学部（2016年度新設）、法政経学部、理学部、工学部、園芸学部、医学部、看護学部の計7学部は、単一学科を含めて計23学科がある。文学部は2016年度から改組されて人文学科のみの単一学科となったが、数十名規模の4つのコースに分かれているため、本論文では人文学科ではなく、4つのコースのそれぞれを便宜的に「学科」と呼ぶ。また、教育学部は5課程に分かれているが、これも「学科」に準ずるものとして「学科」と呼ぶ。薬学部は1年次の段階では学科に分かれていないが、便宜的に学部全体をひとつの「学科」と呼ぶ。千葉大学では学科と課程を総称して「学科等」と呼ぶ習慣もあるが、「学科別」を「学科等別」と表記することによって「学科+等別」という誤解をされることを避けるため、あえて「学科別」と表記することにした。以上の考えに基づき、本論文では千葉大学1年次学生全体を10学部33学科に分類して集計した。受験者数は2,359名で、うち13名は編入生や過年度入学者である。また、聴覚障害のため Section I の受験を免除された学生1名は

本研究に含めなかった。

以下の手順によって個人情報の取扱には十分に留意した。TOEFL ITP の受験番号は数字しか使用できないため、千葉大学の8桁の英数字からなる学生証番号を8桁の数字に変換した形で受験番号が各受験者に付与された。採点後に個人別スコア一覧表がCSV形式で届いてから、Excelを使用して千葉大学の学部学科別の学生名簿と自動照合して、筆者が受験者の氏名を見ることなく学部学科別の集計をした。ただし、受験者が受験番号をマークミスすることによって自動照合に失敗した場合や、同一受験番号のデータが複数存在する場合のみ、採点結果と学生名簿の氏名と生年月日を筆者が確認の上、修正して集計に加えた。学部別学科別集計結果は個人情報とは言えないが、慎重を期して学部名を「A学部」から「J学部」までの英字表記にし、学科名を「学科1」から「学科33」までの数字表記に置き換えて示した。土肥・張（2014）でも当時の9学部に合わせて「A学部」から「I学部」までの英字表記を使用した。本研究の10学部の英字表記とは対応させていない。一般的に統計分析をする際、本来は集団別（学部学科別）の人数を示すべきであるが、学部名や学科名の特定を避けるために、本論文ではあえて示さなかった。なお、もっとも人数が少ない学科の受験者数は20名、多い学科は380名で、1学科あたりの受験者数の平均は71名であった。

#### 4. 結果と分析

以下の4.1～4.5の項目順に「結果と分析」を示したが、平均や標準偏差等の数値は、丸めの誤差を含む。今回の研究では、統計的検定は実施しなかった。

##### 4.1 Totalスコア帯別分布

まず全受験者2,359名のTotalスコアについて、30点刻みのスコア帯別人数を図1に示した。図1の情報は本論文に掲載しただけでなく、各学生が自身の現在の位置付けを把握し、今後の学習目標を設定するための参考となるよう掲示によって学内に周知した。

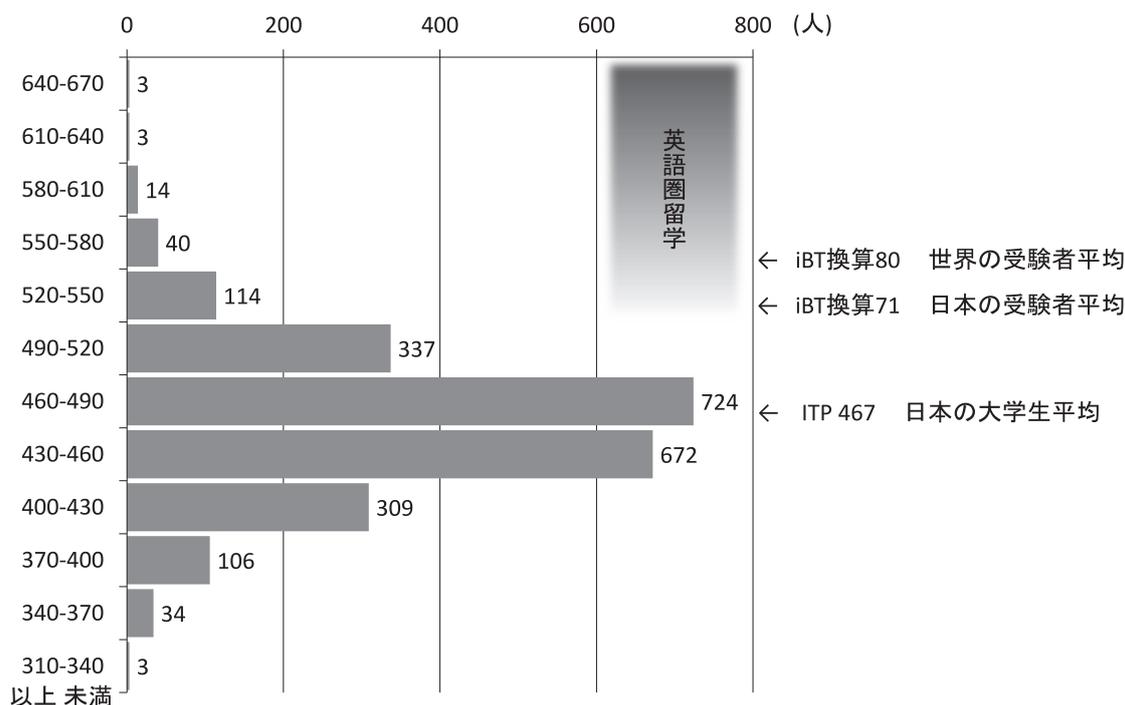


図1 全受験者のTotalスコア帯別分布

2007年から2015年まで9年間にわたって新生に実施してきた TOEIC L&R IP の分布と同様に、ほぼ正規分布に近い形を示し、入試を通過したばかりの新生であっても英語力の差が大きいことが判明した。千葉大学の平均点は461点で、全国の大学生平均（467点）に近い（ただし、TOEFL ITPを実施するのは比較的英語力が高い大学と言われることもある）が、世界の受験者の平均（TOEFL iBT 80）と比べてかなり低い。英語圏への正規留学が可能な英語力の目安とされることの多い550点以上の学生数は60名（2.5%）いたが、逆に言えば大部分の学生の英語力が留学レベルに達していないことを示す。平均点の461点は、近年、基準にされることが多い CEFR（Common European Framework of Reference for Languages）との対応表（国際教育交換協議会2016c）で言えば、A2レベル（Basic User — Waystage: 337～459）とB1レベル（Independent User — Threshold: 460～542）の中間ほどである。より実践的なレベルと言えるB2レベル（Independent User — Vantage: 543～626）に相当するのは69名（2.9%）、C1レベル（Proficient User — Effective Operational Proficiency: 627～677）に相当するのは5名（0.2%）であった。

#### 4.2 学部別Totalスコア平均点

次に、学部別のTotalスコア平均点をA学部からJ学部まで、高い順に図2に示した。500点を超える学部が2つあったが、もっとも高いA学部ともっとも低いJ学部では平均点の差が88点もあり、学部別平均値の比較で見てもばらつきが大きいことが確認できた。

## 千葉大学における TOEFL ITP のスコア分析

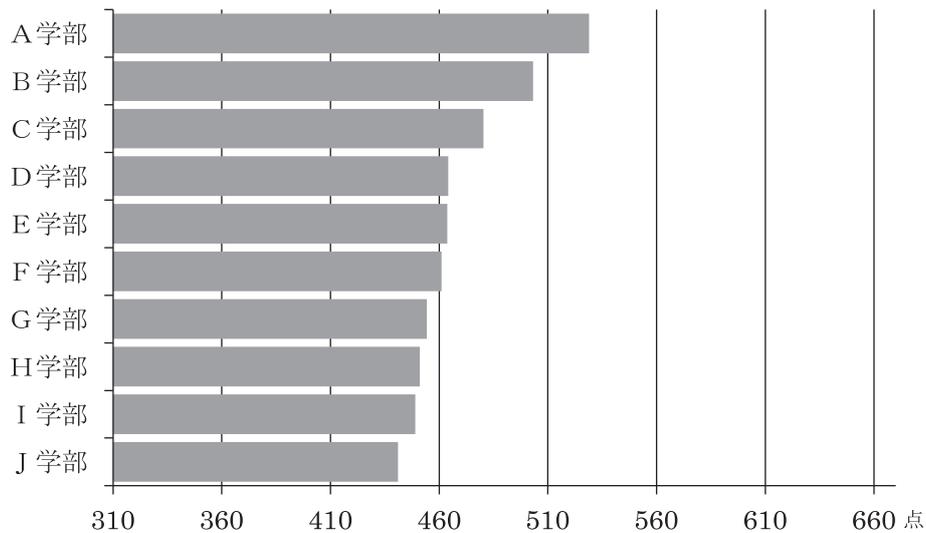


図2 学部別Totalスコア平均点

### 4.3 学科別傾向

より詳細な分析を行うため、千葉大学の33の学科（前述の通り本論文では、学科・課程・コースを総称して「学科」と呼ぶ）に分類してTotalおよびSection別の平均点を表1の左側に示した。学科はTotalの平均点が高い順に「学科1」から「学科33」と名付けた。学科別で見ると、図2の学部別よりも差が広がり、もっとも高い学科と低い学科のTotalの平均点の差が98点もあった。Section I, II, III はそれぞれListening Comprehension、Structure & Written Expression、Reading Comprehension を示し、各Sectionの平均点の差は一見小さく見えるが、これは3つのSectionの平均点を10倍してTotalスコアを求めるといふ計算方法によるものであり、Totalと同様の比較をするのであれば各Sectionの数値を10倍して考える必要がある。たとえばSection Iの平均点をもっとも高い学科と低い学科の差が6.4点しかないが、実質的に64点の差があるという考え方をすべきである。

表1 学科別Section平均と類型

学科名 (数字表記)	Total	Section			高中低			類型						
		I	II	III	I	II	III	低高中	低中高	中高低	中低高	高中低	高低中	
1	529	50.4	54.9	53.4	低	高	中	○						
2	503	48.8	51.3	50.8	低	高	中	○						
3	480	48.2	48.5	47.5	中	高	低			○				
4	477	46.3	48.6	48.2	低	高	中	○						
5	468	46.5	48.2	45.8	中	高	低			○				
6	465	46.6	46.8	46.1	中	高	低			○				
7	465	47.3	46.8	45.5	高	中	低						○	
8	464	46.0	46.4	46.7	低	中	高		○					
9	463	46.4	45.9	46.4	中	低	高					○		
10	463	45.0	47.9	45.9	低	高	中	○						
11	462	46.4	46.6	45.7	中	高	低			○				
12	462	46.2	46.7	45.8	中	高	低			○				
13	460	45.9	46.4	45.6	中	高	低			○				
14	460	45.4	47.1	45.4	中	高	低			○				
15	458	45.5	46.3	45.7	低	高	中	○						
16	458	45.6	45.7	45.9	低	中	高		○					
17	456	46.5	45.4	45.1	高	中	低						○	
18	455	45.6	45.9	44.9	中	高	低			○				
19	455	44.7	45.8	45.9	低	中	高		○					
20	455	45.4	45.9	45.1	中	高	低			○				
21	454	45.9	45.2	45.2	高	中	低						○	
22	453	45.3	45.4	45.3	-	-	-							
23	452	45.1	45.4	45.3	低	高	中	○						
24	452	45.7	45.1	44.8	高	中	低						○	
25	451	45.7	44.7	45.0	高	低	中							○
26	449	44.8	45.0	45.0	低	高	中	○						
27	449	45.5	45.5	43.7	高	中	低						○	
28	446	45.6	44.2	44.1	高	中	低						○	
29	444	45.0	44.5	43.6	高	中	低						○	
30	443	44.6	44.8	43.6	中	高	低			○				
31	439	45.2	42.3	44.3	高	低	中							○
32	437	44.6	43.0	43.4	高	低	中							○
33	431	44.0	43.2	42.0	高	中	低						○	
全学科	461	46.0	46.3	45.9	中	高	低			○				
「全学科」を除く「高」の数					11	17	10							
「全学科」を除く「中」の数					11	11	4							
「全学科」を除く「低」の数					10	4	18							
「全学科」を除く各類型の数								7	3	10	1	8	3	

学科名は Total の平均点が高い順に 1~33の数字表記で示した。

Section別平均点は小数点以下第2位を四捨五入して表記し、同一の場合は第2位以下も比較した。

「学科22」は、Section I と III が45.2857142857143で同一であり、高中低の分類から除いた。

表1の中央の「高中低」欄のIからIIIの3列には、同一学科内でのSection別スコアの相対的比較をした。たとえば「学科1」のSection I、II、IIIの平均点はそれぞれ50.4、54.9、53.4であり、Section IIが「高」、Section IIIが「中」、Section Iが「低」の順となる。Section別の数値には絶対的な基準があるのではなく、世界の受験者平均との相対的な比較をしているに過ぎないが、このような比較をすることによって、学科別の得意分野と苦手分野の傾向を示せると考えた。なお、「学科22」はSection IとSection IIIの平均点が四捨五入する前の段階でまったく同一であり、高中低の区分を示さなかった。今回の研究では統計的検定をしていないが、他の学科についても統計的有意差のない差であっても高中低の区分をつけてしまっている可能性がある。多少の不正確さが含まれていても32の学科すべてを示すことによってある程度の傾向の違いを示すことを目指した。

さらに表1の右の「類型」欄には、「高中低」欄に応じて学科別の型を分類して示した。たとえば、「学科33」はSection I (44.0) が「高」、II (43.2) が「中」、III (42.0) が「低」であるので、「類型」欄の「高中低」型のところに○印を付した。

この結果、「中高低」型の学科数が10となってもっとも多く、「全学科」が「中高低」型に分類されることとも一致した。千葉大学学生の典型的な傾向は、「文法や語法が比較的得意であるが、聴解がやや苦手で、読解力はさらに低い」ということになる。次に多い類型は「高中低」型の8つの学科で、Totalスコアがもっとも低い「学科33」や比較的下位の学科がこの傾向を示している。これはSection I、II、IIIと進むにつれて使用される語彙や概念が難しくなる傾向と一致しているのかもしれない。その次に目立つ類型は7つの学科の「低高中」型であり、Totalスコアが突出して高い「学科1」「学科2」を含む比較的上位の学科が示す傾向である。受験勉強に成功した学生は文法や語法が得意で、読解よりも入試の配点が低い聴解を軽視する傾向があったのかもしれない。この傾向はTOEIC L&R IPの分析（土肥・張2014）の際に、上位の学生群のReadingスコアが（Listeningとの相対的比較において）高かったことと一致する。

その他の3つの類型、つまり「低中高」「中低高」「高低中」は、該当する学科数が比較的少なかった。この3つの類型に共通する点は、(Section Iの聴解は別にして)Section IIよりSection IIIのほうが高いということである。この3種類の学生は、比較的短い文を使用する文法や語法の問題が得意で、語彙や概念の難易度が高い長文読解は苦手ということかもしれない。

以上の類型の比較は、本研究で初めて試みたものであり、筆者の個人的見解を述べただけに過ぎず、仮説レベルのものである。個人差が大きいことは当然の前提であるが、今回の受験者群の傾向だけを見て一般化してしまうことは早計であるかもしれない。今後データの収集と分析を積み重ねることによって、より精緻な解釈を試みたい。

#### 4.4 低得点者の問題

前掲の表1と学科の順は同じであるが、学科別にTotalおよび各Sectionの平均、標準偏

差、最高点、最低点を示したのが表2である。

表2の情報は、2015年度までのTOEIC L&R IPと同様に、学部名と学科名を明記した上で学部教育委員会で各学部の教務委員長にも伝えてあるので、学部学科の単位で活用されているはずであるが、本研究においてとくに注目すべきは、各Sectionの「最低」の列に「31」という数値が目立つことである。各Sectionの理論上の最低点は31点で、31点を

表2 学科別Section別集計

学科名 (数字表記)	Total				Section I (L Comp)				Section II (Str&Wr Exp)				Section III (R Comp)			
	平均	標準 偏差	最高	最低	平均	標準 偏差	最高	最低	平均	標準 偏差	最高	最低	平均	標準 偏差	最高	最低
1	529	39.0	663	453	50.4	4.9	68	42	54.9	4.7	68	43	53.4	4.4	65	40
2	503	39.2	653	417	48.8	4.1	66	37	51.3	5.0	65	37	50.8	5.2	67	35
3	480	33.8	563	397	48.2	3.5	59	41	48.5	3.7	59	39	47.5	5.4	58	31
4	477	25.0	550	433	46.3	2.2	50	41	48.6	3.9	56	41	48.2	4.5	59	37
5	468	36.6	530	370	46.5	2.3	53	41	48.2	5.6	62	34	45.8	5.7	56	34
6	465	40.4	547	337	46.6	3.8	53	31	46.8	5.7	60	31	46.1	5.6	56	31
7	465	30.8	533	403	47.3	2.7	54	42	46.8	4.4	55	37	45.5	5.6	55	31
8	464	36.2	603	347	46.0	3.0	62	32	46.4	5.3	62	31	46.7	5.3	59	31
9	463	30.7	517	357	46.4	2.7	53	39	45.9	4.9	55	31	46.4	4.0	54	37
10	463	33.5	597	397	45.0	3.4	58	37	47.9	4.7	65	39	45.9	4.6	56	34
11	462	32.0	533	357	46.4	2.9	53	39	46.6	4.8	57	31	45.7	5.0	55	31
12	462	25.5	503	367	46.2	3.5	55	37	46.7	3.8	53	36	45.8	3.7	53	37
13	460	29.1	517	400	45.9	2.6	52	41	46.4	4.7	55	31	45.6	4.3	55	37
14	460	44.3	583	383	45.4	4.2	57	34	47.1	5.7	62	31	45.4	5.9	58	35
15	458	34.9	550	347	45.5	3.5	53	32	46.3	4.8	59	31	45.7	5.1	55	31
16	458	38.9	573	383	45.6	3.7	60	37	45.7	4.9	59	31	45.9	5.8	58	31
17	456	33.6	590	373	46.5	3.4	66	39	45.4	4.7	55	31	45.1	4.5	56	31
18	455	37.1	553	373	45.6	3.5	58	37	45.9	4.5	55	34	44.9	5.6	58	31
19	455	28.8	520	367	44.7	3.0	53	34	45.8	4.6	56	31	45.9	4.4	56	31
20	455	38.6	580	357	45.4	3.5	60	37	45.9	4.8	57	31	45.1	5.9	58	31
21	454	35.3	527	350	45.9	2.7	52	39	45.2	5.5	59	31	45.2	4.8	52	31
22	453	24.7	510	410	45.3	2.7	50	37	45.4	4.2	55	34	45.3	5.3	55	31
23	452	32.7	513	357	45.1	3.1	50	39	45.4	4.5	56	36	45.3	5.4	53	31
24	452	30.5	540	370	45.7	2.8	55	39	45.1	4.6	55	34	44.8	4.6	53	31
25	451	30.7	530	387	45.7	2.1	52	41	44.7	5.1	55	31	45.0	4.9	53	34
26	449	37.3	510	350	44.8	3.5	50	34	45.0	5.2	52	31	45.0	4.8	53	31
27	449	42.2	543	363	45.5	4.5	58	32	45.5	5.1	55	31	43.7	6.2	56	31
28	446	38.4	527	347	45.6	2.6	52	39	44.2	5.7	56	31	44.1	5.7	56	31
29	444	37.0	530	343	45.0	4.1	53	32	44.5	4.6	55	31	43.6	5.5	56	31
30	443	40.5	567	357	44.6	4.8	62	32	44.8	5.1	53	34	43.6	5.5	55	31
31	439	32.1	523	383	45.2	3.5	52	39	42.3	4.3	53	34	44.3	4.3	52	35
32	437	32.8	507	337	44.6	2.6	51	32	43.0	4.9	55	31	43.4	5.3	55	31
33	431	40.2	533	330	44.0	3.9	57	31	43.2	5.4	54	31	42.0	5.7	55	31
全学科	461	41.1	663	330	46.0	3.6	68	31	46.3	5.6	68	31	45.9	5.7	67	31

取った学生がいる「学科数」は、表3にまとめたとおり、Section I、II、IIIのそれぞれで2、19、22であった。同一学科で複数の学生が理論上の最低点を取っている可能性があるため、個別データにさかのぼり31点を取った全学の「実人数」を数えると、Section I、II、IIIのそれぞれで2、34、68であった。なお、Totalの理論上の最低点は $(31 + 31 + 31) \times 10 / 3$ の310点であるが、今回の受験者に該当者はいなかった。これは全Sectionで31点を取った学生がいなかったことを示す。

表3 TOEFL ITP 理論上の最低点と最高点の取得者数

	2016学科数	2016実人数
Section I 最低点 (31点)	2	2
Section II 最低点 (31点)	19	34
Section III 最低点 (31点)	22	68
Total 最低点 (310点)	0	0
Section I 最高点 (68点)	1	1
Section II 最高点 (68点)	1	1
Section III 最高点 (67点)	1	1
Total 最高点 (677点)	0	0

Section によって人数が異なる理由は不明であるが、筆者の推測では以下のように考えられる。Section I は音声のペースに合わせて受験するので、時間配分に問題は起きにくく、標準偏差を見てもわかるように差がつきにくい。一方で、Section II と III は各受験者がペース配分を考える必要があり、時間が足りない学生が増えた可能性が高い。とくに、Section III は長文を読む必要があつて難しく、時間が足りない学生が多かつたはずである。また、最初の Section I は音声英語で内容も語彙も比較的易しく解答意欲が保てたが、II や III に進むにつれて難易度が高くなることと上記の「時間不足」の問題も相まって、心理的に疲れ、意欲が低下したり諦めたりした学生が増えた可能性がある。

理由はともかくとして、2007～2015年度の新生が義務として受験したTOEIC L&R IP の各 Section で理論上の最低点の5点を取った学生は、表4に示した通り、過去9年間の合計を9で割ると年平均でListeningが0.6名、Readingが0.8名であった。今回の TOEFL ITP では、本来は英語力に差のあるはずのべ104名もの学生に最低点の31点を出してしまつて、能力の弁別ができていないことになる。最低点を取った学生は本当に英語力が低かつた可能性もあれば、真面目に受験しなかつた可能性もあるが、TOEIC L&R IP 実施の際には顕在化していなかつたTOEFL ITP 実施の問題点として指摘したい。留学のために英語力の比較的高い上位層、中位層の学生だけを弁別することが目的であればこれでよ

いのかもしれないが、下位層の学生は試験中に「難しかった」という否定的な印象を持ち、とくに最低点を取ってしまった学生が結果を通知された際に「あなたのスコアは、千葉大学でも日本でも世界でも最低の英語力を示します」という誤った烙印を押すことになりかねず、テスト結果の解釈だけにとどまらず、その後の英語嫌いを助長する心配もある。これは、テストそのものの問題ではなく、あくまでもテストを選択する受験者や団体との組み合わせの問題である。

表4 TOEIC L&R IP 理論上の最低点と最高点の取得者数

		2007	2008	2009	2010	2011	2012	2013	2014	2015	9年間の計	年平均
最低点	Listening (5点)	0	0	0	5	0	0	0	0	0	5	0.6
	Reading (5点)	0	1	1	2	0	1	1	1	0	7	0.8
	Total (10点)	0	0	0	0	0	0	0	0	0	0	0.0
最高点	Listening (495点)	4	14	7	12	10	2	10	6	6	71	7.9
	Reading (495点)	0	0	1	1	0	1	1	2	0	6	0.7
	Total (990点)	0	0	1	1	0	0	1	2	0	5	0.6

念のために上記とは逆のケース、つまり理論上の最高点を取った人数を表3と表4で比較してみると、TOEFL ITP で理論上の最高点（67または68点）を取った学生は Section I, II, III の合計でのべ3名、Total で理論上の最高点（677点）を取った学生がゼロであったのに対し、TOEIC L&R IP で理論上の最高点（495点）を取ったのは Listening, Reading のそれぞれで1年あたり7.9名と0.7名で、Total で理論上の最高点（990点）を取った学生は0.6名であった。TOEIC L&R IP の Reading よりも Listening のほうが該当者数が多いのは、土肥・張（2014）でも論じたように TOEIC L&R IP は Listening の平均点のほうが高く出る特性があることが影響していると考えられる。学生の中に英語の母語話者や海外経験が長い者も含まれる可能性を考えると、非母語話者の英語力を測定することが目的の TOEFL ITP や TOEIC L&R IP で得点が飽和することはある程度やむを得ないことであるし、最高点を取ったことによって学生を落胆させるようなことは少なく、悪影響も少ないと言えるであろう。

#### 4.5 TOEIC L&R IP との換算

土肥・張（2014）はTOEIC L&R IP TotalとTOEFL ITP Totalの既存の換算式

$$[\text{TOEFL ITP Total}] = [\text{TOEIC L\&R IP Total}] \times 0.348 + 296 \quad (\text{旧換算式})$$

の問題点を指摘し、新しい換算式の作成を試みた。具体的には、過去10年間でTOEIC L&R IPを受験したのべ21,280名の受験者のうち、2か月以内の間隔でTOEFL ITPも受験した138名について両テストのスコアの単回帰分析を行い、

$$[\text{TOEFL ITP Total}] = [\text{TOEIC L\&R IP Total}] \times 0.229 + 344 \quad (\text{新換算式})$$

という換算式を作成した。その後、この換算式が学内で採用され、2016年度以降について

外部テストを使用した単位認定（表5）や授業履修基準（表6）に際して、下記のような基準変更を行った。

表5 2016年度以降の「検定英語」単位認定のスコア基準

旧「検定英語I・II」	(計4単位)	→	新「検定英語」(4単位)
TOEFL ITP, PBT	550点以上	→	<u>510点以上</u>
TOEFL iBT	79点以上	→	<u>64点以上</u>
TOEIC (IP)	730点以上	→	730点以上 (変更なし)
IELTS	6以上	→	6以上 (変更なし)
英検	1級	→	1級 (変更なし)

表6 2016年度以降の英語III・英語IV・英語V・英語VI の履修基準

新科目名 (旧科目名)	TOEFL ITP TOEFL PBT	TOEFL iBT	TOEIC L&R 公開 TOEIC L&R IP	IELTS
英語III (中級英語I)	487 → 470	57 → 52	550 (変更なし)	5 (変更なし)
英語IV (中級英語II)	523 → 493	69 → 58	650 (変更なし)	5.5 (変更なし)
英語V (上級英語)	550 → 510	79 → 64	730 (変更なし)	6 (変更なし)
英語VI (上級英語)				

本研究では、今年度の TOEFL ITP の結果に基づいて試算し、新換算式の妥当性と信頼性を検証してみることにした。今回の TOEFL ITP 全学平均が461点で、旧換算式と新換算式で TOEIC L&R IP に換算すると、それぞれ 474点、510点である。過去の新生の英語力と2016年度の新入生の英語力を比較する明確なデータは存在しないが、2011～2015年の新生の TOEIC L&R IP 平均点はそれぞれ495、501、509、491、499点で、500点前後で安定していた。2016年度の新入生は TOEIC L&R IP を受験していないので同条件での比較はできないが、新設の国際教養学部の TOEFL ITP の平均点が全学の平均点より高いという事実を考慮すると、2016年の新生全体の英語力は若干高めであるのかもしれない。この推定が正しければ、2016年度の新入生の英語力が旧換算式に基づく TOEIC L&R IP換算で474点と大きく下落したとは考えにくく、新換算式に基づく510点という若干高めの数字になったことは、ほぼ妥当な結果と言えそうである。

2点目の検証結果は、英語III以上（旧中級英語I以上）の基準による換算である。2015

年度新入生の TOEIC L&R IP の結果に基づく試算では、764名が旧中級英語I以上を履修できる英語力があつた。2016 年度新入生の TOEFL ITP の結果については、旧換算式による487点以上が基準であるとしたら581名で人数が24%も減少してしまう。新換算式による470点以上が基準であるとしたら978名となって人数が28%増となる。上記と同じ理由で、減るよりは増えるほうが妥当と推定している。

その一方で、新換算式に若干の疑問が生じるデータも2つある。そのひとつ目は TOEFL ITPの平均点をもっとも高かつたA学部の試算結果で、TOEFL ITP平均点、旧換算式に基づく TOEIC L&R IP、新換算式に基づく TOEIC L&R IP のそれぞれが、529、669、808であつた。2016年は英語力の高い学生が入学したという可能性もあるが、新換算式で TOEIC L&R IP が 808点になってしまうのは、2015年のA学部の TOEIC L&R IP の平均点が688であつたことと比べて高すぎるような印象を受ける。

もうひとつ新換算式の問題点と考えられるデータがある。それは検定英語の4単位に相当する TOEIC L&R IP 730点の換算結果で、旧換算式と新換算式ではそれぞれ TOEFL ITP の550点と510点に相当する。2015年4月の TOEIC L&R IP で検定英語の4単位に相当する730点以上を取得したのが130名であつたのに対し、2016年4月の新入生の TOEFL ITPの結果では、旧換算式による550点以上取得者が60名で、新換算式による510点以上取得者が252名であつた。60名と半減以下になるのは少なすぎる半面で、倍近く252名になるのも多すぎる印象を受ける。2016年の新入生のほうが若干英語力が高そうだと推定しても、それだけで説明できる増え方ではなさそうである。

上記2点の問題点について、もし(とくに上位群で)新換算式が信頼できないものであるとしたら、その原因は不明であるが、ひとつの可能性は土肥・張(2014)の研究で中間層のデータが比較的多く、高得点者(および低得点者)のデータ数が多くなかつたという問題があるかもしれない。さらに言えば、数百名以上の換算をするための根拠となる基礎データが138対だけでは不十分であつたのかもしれない。いずれにせよ、今後、より多くのデータを蓄積できれば、多かれ少なかれ換算式は変わるであろう。しかし、新しく採用したばかりの換算式をすぐに変えてしまつては教育現場が混乱するし、もし旧換算式も新換算式も問題があるとしても、当面の間はこれ以上のデータがないので再度の改訂は不可能である。

## 5. 研究のまとめと今後の課題

本研究は、2016年度から千葉大学の新生を対象に実施することとなつた TOEFL ITP について、2016年4月の1回分、2,359名への実施結果について分析したものである。得点の分布、学部別平均点の比較、学科別平均点の比較のいずれを見てもばらつきが大きいことが判明した。全学の平均点は461点で全国の大学生平均に近い数字であつたが、CEFRのB2レベル以上に相当する543点以上が3.1%で、正規留学等が可能な実用レベル

に達している学生は少ないと結論した。3つのSectionの比較をしてみると、全学的に同じ傾向を示すのではなく、学科別の傾向が異なることが判明した。また、TOEIC L&R IP を全学的に実施した場合と比較して、TOEFL ITP ではSection別の理論上の最低点を取る学生が多いことを問題点として指摘した。最後に、土肥・張（2014）が提案した TOEFL ITP と TOEIC L&R IP の換算式について、今回のデータと過去のデータも比較しながら多面的に検証し、妥当性と問題点について論じた。

筆者らが過去に行った TOEIC L&R IP についての3研究では、土肥（2006）が自主的に受験した2年分のべ1,240名、土肥・柳瀬（2009）が全学の新入生が受験した1年分を含む5年分のべ6,689名、土肥・張（2014）が全学の新入生が受験した6年分を含む10年分21,280名という順に規模を段階的に拡大し、継続的かつ包括的で信頼性の高い研究を目指した。本論文の TOEFL ITP についての研究は、2016年4月に実施した1回分のみ2,359名の結果分析であり、複数回受験者のデータは含まない。土肥・張（2014）は、少数ながら6年間で TOEFL ITP を複数回受験した80名について、平均381日の間隔で18点上昇したとの結果を示した。今後 TOEFL ITP 受験者が大きく増えることによって、より信頼性の高いデータが得られると考えている。

本研究は入学時のテスト結果のみに焦点を絞ったが、大学における教育方法の改善がより重要であることは言うまでもない。しかし、千葉大学では長年にわたって英語教員数の削減が続けられ、学生の卒業要件単位も削減されるなどして、見通しは明るくない。教員ひとりひとりの努力や工夫には限界があり、より適切な教育環境が整うことによって学生の学習効果につながることを願いたい。

## 謝辞

本研究の実施に際して、千葉大学教職員や学生のご協力を得た。また、国際教育交換協議会（CIEE）日本代表部から各種資料のご提供をいただくとともに、本論文の草稿に対して専門的見地から貴重なご助言をいただいた。さらに、査読者からもご意見をいただいて本論文の改善に資することができた。ここに感謝の意を表したい。

## 主な参考文献

- 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.
- 土肥充、張智君、「千葉大学における TOEIC IPと TOEFL ITPのスコア分析と経年調査」、『言語文化論叢』、千葉大学言語教育センター、第8号、pp.15-32、<http://f.chiba-u.jp/about/plc08/plc08-02.pdf>, 2014.
- 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](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/](http://www.ets.org/toeic/speaking_writing/about/), 2013d.
- 国際教育交換協議会 (CIEE) 日本代表部、「TOEFL ITPテスト概要」2016年4月、2016a.
- 国際教育交換協議会 (CIEE) 日本代表部、「TOEFL ITP: テストの構成」、<http://www.cieej.or.jp/toefl/itp/composition.html>, 2016b.
- 国際教育交換協議会 (CIEE) 日本代表部、「TOEFL ITP: 測定するレベル」、<http://www.cieej.or.jp/toefl/itp/level.html>, 2016c.
- 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](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プログラムDATA & ANALYSIS 2016」、2016.
- 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.