



入学時英語能力とその推移に関する一考察

田中, 順子

(Citation)

国際文化学研究 : 神戸大学国際文化学部紀要, 18:1*-20*

(Issue Date)

2002-09

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCOI)

<https://doi.org/10.24546/81001252>

(URL)

<https://hdl.handle.net/20.500.14094/81001252>



入学時英語能力とその推移に関する一考察

田中 順子

1. はじめに

神戸大学国際文化学部は、「国際的な理解と協調という目的のために、異文化理解を促進し、それを支えるコミュニケーションの方法を解明するための教育研究を行うこと」を理念として設置された（神戸大学国際文化学部自己評価委員会，2001，pp.1-2）。その設置理念に基づき「グローバル化時代に対応できる〈活動する知性〉を身につけた人材の育成」（pp.1-2）を目指して、「外国語運用能力の向上」，「異文化理解能力の育成」，「情報処理能力の開発」の3点を柱とした教育目標を掲げている（p.2）。教育目標として「外国語運用能力の向上」を掲げているように，本学部では高い運用能力の獲得に結びつくような外国語教育を提供することが望まれる。

特に，国際コミュニケーションの媒体として使用機会が多い英語については，高い能力の開発が様々な方面から求められている。経済・産業界では実践的な英語能力獲得の必要性が一層高まってきており，英語能力が実務に必要なばかりか，昇進や昇格に必要な具体的な英語能力レベルを設定している企業もある（池野，2002，pp. 48-51）。一方教育分野では，英語圏への留学を希望する場合，英語で授業履修や学習が行える英語能力レベルにあることを証明するよう志望校から要求される。このように，信頼性の高い英語能力評価基準に照らし合わせて，実践的な英語能力レベルを客観的に証明する必要性が高まっている。経済・産業界では“Test of English as International Communication”（TOEIC）が¹，英語圏（特に北アメリカ）への留学希望者の場合は“Test of English as a Foreign Language”（TOEFL）が英語能力の客観的な指標を出すための代表的な手段として多くの者に選ばれている²。

上で述べたように「外国語運用能力の向上」という本学部の教育目標の達成状況を，広く認知された英語能力試験を用いて計測することは，近年の社

会的風潮や要請から言っても妥当である³。また、社会や学生のニーズに応えるような教育を提供していくためには、学生の英語学習の現状や特色について実態調査を実施して基礎資料を集め、学生の要望を適切に把握しておくことが必須である。この二つの観点から、本学部一年生の英語能力を世界的に認知された英語能力指標に基づいて計測し、彼らの中に見られる様々な特質が英語学習にどのように関わっているのかを探るために「国際文化学部新入生の英語学習動機の分析と英語能力養成への展望」を実施した⁴。本稿ではこの研究の一部を報告する。

2. 研究の目的

本稿で報告する研究の主な目的は、上でも述べたように、国際的に認知された信頼性の高い試験で本学部在校生の英語能力を計測し、計測したスコアをもとに考察を加えることである。具体的には本学部に入学者の英語能力を、前期授業開始近くと後期授業終了近くの2回計測し、1回目と2回目の間の8ヶ月間にスコアがどう推移したかを分析する。大学受験から時間が経つほど英語能力は低下するのではないかと仮説されるが、実際にこの仮説が正しいのかを検証する。副次的な目的としては、将来のカリキュラム評価の基礎となるベースラインデータを提供することである。

3. Test of English as a Foreign Language (TOEFL)の成り立ちと有用性

本研究では英語能力を測るために使用するテストとしてTOEFLを選択した。TOEFLはEducational Testing Service(本部アメリカ合衆国ニュージャージー州、以下ETSと略す)が実施している標準検査(standardized test)である。世界210カ国以上で毎年80万人以上が受験し、4200以上の教育機関や政府機関が利用しており(Ewing, 2002)、その影響力は世界的規模で増大している(Pierce, 1992)。この研究では、TOEFLのなかでも“*Institutional program, Test of English as a Foreign Language*”(ITP TOEFL)と

呼ばれる TOEFL の団体特別受験制度を選択した。国際的に認知されている英語能力試験としては TOEFL 以外に TOEIC も考えられる。しかし、本学部生の中には留学希望者が比較的多く、英語圏に留学するためには TOEFL を近い将来に受験する必要があるため、1 年次から TOEFL の設問形態に慣れておくことは有益だと考えて TOEFL を選択した。

本研究で用いた英語能力試験の特色や有用性についてここでさらに説明を加えておく。TOEFL の出現以前は、多様な英語能力の側面を測ることができる客観的な標準英語能力試験で、かつ全米規模で実施されるような試験は存在していなかった。1961年にそのような試験の必要性の確認と開発への合意がなされ(Spolsky, 1990)、テスト開発のために米国政府機関や私立団体の代表者からなる協議会が1962年に設立され、この協議会が1963-64年度から TOEFL の開発を支援した(Educational Testing Service, 2001, p. 2)。TOEFL は当初 Modern Language Association によって実施されたが、1965年からは College Board と ETS が共同で実施し、1973年以降は College Board と Grade Record Examinations の委託を受けて ETS が TOEFL の実施団体となっている (Educational Testing Service, 2001, p. 2)。

TOEFL のテストフォームが現在のように3つのセクションに分けられたのは1976年である。これ以降は3つのセクションに分かれた設問形態が取り入れられている(Educational Testing Service, 2001)。1995年7月に設問形態や設問数が改編され(Educational Testing Service, 2002)、現在使用されている紙媒体の試験形態へと至っている。現行の3つのセクションは、リスニング(listening comprehension)、構造(structure and written expression)、読解(vocabulary and reading comprehension)に分けられている。その再試験信頼性は以下の表1に示されるように、リスニングで.87、構造で.81、読解で.88、3セクションを合わせたテスト全体では.94と報告されており(Educational Testing Service, 1998)、非常に高いといえる⁵。これら3つのセクションから成る TOEFL に加えて、英語の書く力を測るテスト(Test of Written English)や話す力を測るテスト(Test of

表 1. TOEFL の 3 セクション別と全体での信頼性および測定の標準誤差
Reliabilities and Standard Errors of Measurement (SEM)

	Section Reliability	SEM
Listening	0.87	1.6
Structure/Writing	0.81	2.1
Reading	0.88	1.6
Total	0.94	10.4

出典：Educational Testing Service. (1998). *Computer-Based TOEFL Score User Guide* (p.34).

Spoken English)が TOEFL を補足する試験として存在している⁶。

現在ETS が行っている TOEFL は、スコアを受験者本人や実施団体以外の教育機関に公式に通知可能かどうかで 2 種類に分けられる。大学への入学者選抜資料にするなどの目的でスコアの公式通知を依頼できる TOEFL (以下、通常の TOEFL と呼ぶ) と、受験者本人と実施団体のみにスコアが通知される ITP TOEFL がある。通常の TOEFL は、紙媒体のテストフォームに記載された問題を解き、解答をマークシートに書き込む Paper-Based TOEFL (TOEFL-PBT)と、コンピュータ上で試験問題の提示と解答を行う Computer-Based TOEFL (TOEFL-CBT)がある。TOEFL-CBT のスコアの最低値は 40、最高値は 300 となっている。

ITP TOEFL は紙媒体のみで行われており、レベル別に 2 つに分類される。一つは、通常の TOEFL と変わらない TOEFL でレベル 1 と呼ばれるものである。これは通常の紙媒体の TOEFL と同様にスコアの最高値が 677 に設定されている。ただし、現行の紙媒体テストと異なり、Test of Written English は入っていない。もう一つの ITP TOEFL は、Pre-TOEFL あるいはレベル 2 と呼ばれるもので、計測できる最高スコアが 500 までの制限がある。

本研究で用いたテストはレベル 1 の ITP TOEFL であり、難易度は通常の紙媒体の TOEFL と同じものである。このテストは、ETS が過去に行った TOEFL で使用したテストフォームの中から一つを使って大学などの団体

が主催して行うものであり、TOEFLと同様にETS（あるいはETSから委任された団体）が管轄している。大学などの実施団体は試験の日時や場所を選択することは出来るが、使用するテストフォームの選択や採点には関与できない。また、試験実施も紙媒体のTOEFLと同様に厳格な監督の下で行われる。このような背景からレベル1のITP TOEFLは、紙媒体のTOEFLの持つ信頼性と同等の再試験信頼性を有していると判断できる⁷。

ETSが世界中の受験者の成績に基づいて発表したTOEFLスコアの統計結果を表2と表3に示す。表2は2000年7月から2001年6月にかけて行われた紙媒体でのTOEFLのセクション別のスコア(受験者数230,877人)を示したものである⁸。これらのスコアは素点ではなく尺度化されたものである。すべてのセクションにおいて最低スコアは31で、最高スコアは68である。ただしリーディングの最高スコアは67となっている。3つのセクションを合わせた総合での最低スコアは310で最高は677である。総合でのスコアの算出方法は、各セクションの尺度化されたスコアの和を10倍して3で割る方法をとる⁹。たとえばセクション1, 2, 3の尺度化されたスコアがそれぞれ40, 50, 60だった場合は、 $(40+50+60) \times 10 \div 3 = 500$ となり、500が総合スコアとなる。ITP TOEFLでも通常のTOEFLと同様の最低スコアと最高スコアの設定になっている。

表2. 紙媒体 TOEFL のセクション別と全体での最低と最高点

Section	Min.	Max.
1. Listening Comprehension	31	68
2. Structure and Written Expression	31	68
3. Reading Comprehension	31	67
Total	310	677

出典：Educational Testing Service. (2001). *TOEFL Test and Score Data Summary* (p.11).

2000年から2001年に行われた紙媒体のTOEFLで、大学学部への留学を希

望する受験者（総数48,893人）が記録したスコアとその百分位数（パーセンタイル値）を表3に示す¹⁰。この表を参照すると個人のスコアが全体の受験者中でどのあたりに位置するのかを推測することができる。

表3. 英語圏の大学学部への留学志望者のセクション別および総合の TOEFL スコアと百分位数

Scaled Score	Section 1 Listening Comprehension	Section 2 Structure and Written Comprehension	Section 3 Reading Comprehension	Total Scaled Score	Percentile Rank
68		99			
66	98	96	99	660	99
64	96	93	98	640	98
62	93	89	95	620	94
60	88	83	91	600	90
58	83	75	84	580	83
56	76	66	75	560	74
54	67	56	64	540	64
52	57	45	52	520	52
50	45	36	41	500	39
48	34	26	30	480	28
46	22	18	21	460	18
44	14	12	14	440	11
42	8	7	9	420	6
40	4	4	5	400	3
38	3	2	3	380	1
36	1	1	2	360	
34	1	1	1	340	
32			1	320	
Mean	51.2	52.8	51.3	Mean	518
S.D.	6.8	7.3	6.7	S.D.	63

“TOEFL test and score data summary” (Educational Testing Service, 2001, p. 12)より転載。

ちなみに、2000年から2001年に紙媒体の TOEFL を受験した者のうち、日本語が母国語である者 34,645 人(学部への留学志望者に限定されない)のセッション別および総合スコアの平均値は、セッション 1 が 49、セッション 2 が 52、セッション 3 が 50、総合が 505 であった (Educational Testing Service, 2001, p.14)¹¹。一例として、総合スコアで 500 を獲得した人の成績の解釈の仕方をあげると、500 というスコアの獲得者は受験者の中で下から約 39 パーセントのところに位置しており、受験者総数の約 39% の人がこれより下のスコアを取り、約 61% の人がこれより高いスコアを取っていることになる。

以上の点をまとめると、TOEFL を英語能力の測定に用いる利点は 2 点ある。第一に、再試験信頼性が高いことである。同じ能力を持った人が試験を受けた場合に、毎回ほぼ同様のスコアを計測することが出来ることになる。そのため、TOEFL でスコアの変動が見られた場合はテスト問題の設計に起因しない要因によってその変動が起こされたと考えることができる。第 2 の利点は、大人数の受験者のデータをもとにスコアが尺度化されており、対応する百分位数も発表されているため、TOEFL 受験者全体のなかで個人の能力がどの位置にあるのかを容易に知ることが出来る点である。通常の TOEFL に見られるこのような特徴は ITP TOEFL にも該当する。これらの 2 点の有用性に加えて、希望の日時に一括して学内で試験が行える利便性から ITP TOEFL を採用することにした。

4. 調査・分析の手続き

4.1 被験者

本研究の被験者は 2001 年度に国際文化学部に入学者 1 年生 144 名のうちの 128 名である。被験者 128 名は全て大学教育研究センターが提供する 1 年生向け英語授業の受講者である。国際文化学部 1 年生の英語授業は 3 クラスに分けて行われており、それぞれ成員は 48 名ずつであった。各回の試験に先立って、各クラスの担当教員が本研究の内容の説明書兼参加への同意書を授業内、

授業外で配布し回収した。各クラスの担当教員には本研究の研究者は含まれていない。被験者は自らの自由意志で本研究に参加して、一回目の TOEFL (TOEFL1)や二回目の TOEFL(TOEFL2)を受験した。本研究への参加と引き替えに学生に報酬が与えられることはなかった。

TOEFL1 を受験した者は123名、TOEFL2 を受験した者は73名で、TOEFL1 と TOEFL2 の2回とも受験した者は68名であった。本稿では、一回目と二回目の結果を比較する必要性から、一回目と二回目ともに受験した68名のスコアに焦点をあてる。この68名の平均年齢は18.65 ($SD = .49$, $Min = 18$, $Max = 20$)歳で、男性が18名、女性が50名であった。下に示す統計データは特記しない限りこの68名についての分析結果である。

4. 2 データ収集と分析

入学時と8ヶ月経過後の英語能力の推移をみるために、ITP TOEFL を使って2度英語能力試験を行った。一回目を前期授業開始後の2001年5月7日に、二回目を後期授業終了前の2002年1月11日に実施した。二回のTOEFLでは異なるテストブックレット(問題)が使用された。二回目実施時に学力自己診断や学習動機の調査を兼ねたアンケートを配布した。

本稿ではTOEFLスコアの量的分析のみを行い、アンケートデータについては言及しない。TOEFLスコアの分析方法としては、(1)二回とも受験した者のスコアについて t 検定を用いて一回目と二回目のスコアの差を吟味し、(2)一回目のスコアと二回目のスコア間のピアソン積率相関を、TOEFLの再検査信頼性の値と比較する。量的データの解析のために用いた統計パッケージはSPSS (version 6.1)である。

5. 結果

5. 1 TOEFL1 と TOEFL2 スコアの比較

TOEFL1 と TOEFL2 の2つのテストについて、それぞれのテストを一回だけ受けた者、両方を受けた者の3つの受験者群に分けて記述統計の結果を

表4, 表5に示した. TOEFL1を受験した123名の結果は, 平均値が471.99, 標準偏差が29.84, 最大値が580, 最小値が403であった. TOEFL1のみを受験した55名の結果は, 平均値が468.56, 標準偏差が25.83, 最大値が537, 最小値が407であった. TOEFL1と2の両方受験した者68名のスコアは平均が474.76, 標準偏差が32.65, 最大値が580, 最小値が403であった.

表4. TOEFL1の結果

	TOEFL1 を 受験した者全員の TOEFL1 スコア	TOEFL1 のみ 受験した者の TOEFL1 スコア	TOEFL1 と 2 を両 方受験したものの TOEFL1 スコア
Sample size (<i>n</i>)	123	55	68
Mean	471.99	468.56	474.76
SD	29.84	25.83	32.65
Max.	580	537	580
Min.	403	407	403

TOEFL1 だけ受験した者($n=55$)の TOEFL1 スコアと, TOEFL1 と 2 の両方を受験した者($n=68$)の TOEFL1 スコアとの差は 8.2 ポイントであり, それほど大きくはない. この 2 つの群間の TOEFL1 スコアの差が統計的に有意な差であるのかどうかに関しては次節で述べる.

表5. TOEFL2の結果

	TOEFL2 を 受験した者全員の TOEFL2 スコア	TOEFL2 のみ 受験したものの TOEFL2 スコア	TOEFL1 と 2 を両 方受験したものの TOEFL2 スコア
Sample size (<i>n</i>)	73	5	68
Mean	479.90	488.80	479.25
SD	26.99	22.74	27.30
Max.	537	513	537
Min.	397	457	397

次に TOEFL1 と TOEFL2 の総合スコアの分布表を表6に示す. TOEFL1 で一番多い得点段階は 461-480 の範囲であり, それに 441-460, 481-500 の範囲と続く. TOEFL2 で一番多い得点段階は 461-480 であり, 441-460 の

範囲と 501-520 の範囲がそれに続く。ここで、TOEFL1 と 2 を受けた68名を、TOEFL1 での得点順で、下位得点層（下位層）、中位得点層（中位層）、上位得点層（上位層）の3つの得点層に分割した。381から460の範囲の得点者（23名）が下位、461から480の範囲の得点者（23名）が中位、481から580の範囲の得点者（22名）を上位とした。表7ではそれぞれの得点範囲の境界を破線で示してある。

TOEFL1 では460以下のスコアをとった者が23人(33.8%)であったのだが、TOEFL2 では18人(26.5%)に減少している。461から480の間のスコアをとったものがTOEFL1 では23人(33.9%)であるのに対してTOEFL2 では26人(38.3%)に増加している。また481から580の間のスコアをとったものは、TOEFL1 では22人(33.4%)であり、それがTOEFL2 では24人(35.3%)と微増している。

表 6. TOEFL1 と TOEFL2 の総合スコア分布表

得点範囲	TOEFL 1			TOEFL 2		
	度数(人)	相対度数(%)	累積相対 度数(%)	度数(人)	相対度数(%)	累積相対 度数(%)
561-580	1	1.5	100.0	0	0	100.0
541-560	2	3.0	98.5	0	0	100.0
521-540	5	7.4	95.6	5	7.4	100.0
501-520	2	3.0	88.2	11	16.1	92.6
481-500	12	17.6	85.3	8	11.7	76.5
461-480	23	33.9	67.6	26	38.3	64.7
441-460	14	17.5	33.8	15	22.2	26.5
421-440	6	8.8	13.2	1	1.5	4.4
401-420	3	4.4	4.4	1	1.5	2.9
381-400	0	0	0	1	1.5	1.5
合 計	68			68		

一回目の試験から約8ヶ月後、460点以下の下位層が減少し、461から480の範囲の中位層が増加し、481から580の範囲の上位層が微増した。下位層が減少し中位層が増加したことは望ましい。しかし、その一方でTOEFL1では541から580の範囲の「高スコア取得者」が3名いたのが、TOEFL2では皆無となっている。入学当初にTOEFL541以上だった高スコア取得者は、TOEFLスコアから判断する限り、この約8ヶ月間の英語授業にも関わらず、英語能力の向上が見られなかったと言える。

次に、リスニング、構造、読解の三つのセクション別の成績を下記の表7に示す。セクション別スコアの百分位数は表3中の百分位数に基づいた概算値である。

表7. TOEFL一回目、二回目のセクション別および全体の平均値と百分位数と相関

	Section 1		Section 2.		Section 3.		Total Scaled Score	
	Listening Comprehension		Structure & Written Comprehension		Reading Comprehension			
	スコア	百分位数 (概算)	スコア	百分位数 (概算)	スコア	百分位数 (概算)	スコア	百分位数 (概算)
TOEFL1	43.35	12.05 [†]	49.44	33.20 [†]	49.65	39.08 [†]	474.76	25.38 [†]
TOEFL2	45.13	18.52 [†]	49.65	36.81 [†]	48.46	32.53 [†]	479.25	27.63 [†]
TOEFL1 と TOEFL2 の 相関係数(<i>r</i>)	.50*		.52*		.46*		.67*	

[†] 上記百分位数は、表3の値を参考に割り出した概算値であって実際の百分位数ではない。

* $p < .001$.

一回目と二回目を比べると、リスニングが1.78ポイント増加し、構造もわずかではあるが0.21ポイント増加している。それに比べ、読解はスコアが1.19ポイント減少している。入学時と入学8ヶ月後とを比べると、リスニングが少し向上し、構造は入学時の力を保持したのに対して、読解の力が減少している。

次に、TOEFL1のスコアに基づいて受験者68名を下位(441-460)、中位(461-480)、上位(481-500)に分け、TOEFL1からTOEFL2間のスコアの伸びの平均を、3セクション別と3セクション総合で示したのが下の表8である。各セクション別でも3セクション総合でも、スコアの伸びは下位層>中位層>上位層の順に大きかった。

表 8. TOEFL1 から TOEFL2 間の得点層別スコア変化の平均

	総合	リスニング	構造	読解
上位層 ($n = 22$)	-9.86	0	-0.59	-2.41
中位層 ($n = 23$)	2.30	1.35	0.83	-1.52
下位層 ($n = 23$)	20.39	3.91	1.91	0.30
全員 ($n = 68$) [†]	4.48	1.78	0.74	-1.19

[†]TOEFL1とTOEFL2の両方を受験した者全員を指す。

5. 2 推測統計の結果

5. 2. 1 TOEFL1のみ受験した群とTOEFL1と2を受験した群の等質性の検討

次に、神戸大学入学後の8ヶ月間に新入生の英語能力にどのような変化がみられたのかを、言い換えればTOEFL1とTOEFL2の間で成績の伸びの関係がどうであったのかを検討した。以下の分析において有意水準はすべて5%に設定した。まず、TOEFL1だけ受験しTOEFL2は受験しなかった群($n = 55$)と、TOEFL1と2の両方を受験した群($n = 68$)が等質な母集団から来た群であったのかどうかを検討するために、TOEFL1の総合点について t 検定を行った。その結果、 $t(121) = -1.15$ ($p = .25$)であり、信頼区間内であるため、TOEFL1だけ受験した群とTOEFL1と2の両方を受験した群の母集団の平均の差はないという仮説は棄却することができず、よってこの2つの群は等質であると結論された。これにより、TOEFL1とTOEFL2間の英語能力の伸びに関する検定はこの両方を受験した群である68名のデータを使って分析を行うこととした。

5. 2. 2 TOEFL1 と TOEFL2 のスコア間の差の吟味

TOEFL1 のスコアと TOEFL2 のスコア間の差が有意なものであるかを検定するために、両方のテストを受験した68名のデータについてその差を吟味した(表7のスコアを参照)。TOEFL1 と TOEFL2 のスコアをリスニング、構造、読解のそれぞれのセクションごとと、テスト全体(3セクション総合)で両側検定で t 検定を行ったところ、リスニングでは $t(67) = -3.56$ ($p < .001$)、構造では $t = -1.52$ ($p = .13$)、読解では $t = 2.42$ ($p < .020$)、テスト全体では $t = -1.50$ ($p = .14$)の結果をえた。

従って、3セクション総合や構造では有意な差がないものの、リスニングと読解セクションのスコアにおいて TOEFL1 と TOEFL2 間で有意な差が見られた。リスニングでは TOEFL2 のスコアが TOEFL1 のスコアよりも有意に高く、読解では TOEFL2 のスコアが TOEFL1 のスコアよりも有意に低かった。この結果から、TOEFL1 よりも TOEFL2 のスコアが低下するという仮説は、読解についてのみ支持された。

5. 2. 3 TOEFL1 スコアと TOEFL2 スコア間の相関

続いて、TOEFL1 と TOEFL2 スコア間のピアソンの積率相関係数を算出した。上記の表7に、全体のスコア間の相関係数とともに各セクションのスコア間の相関係数を示してある。TOEFL1 と TOEFL2 間で、リスニングスコア間の相関係数が $r = .50$ ($p < .001$)、構造のスコア間では $r = .52$ ($p < .001$)、読解のスコア間では $r = .46$ ($p < .001$)であった。TOEFL の再試験信頼性は、総合で $\alpha = .94$ 、リスニングが $\alpha = .87$ 、構造が $\alpha = .81$ 、読解が $\alpha = .88$ であった(Educational Testing Service, 1998)ことから、本研究で得られた相関係数は TOEFL の高い信頼性から比すると比較的低かったと判断される。このことは逆に TOEFL1 と TOEFL2 との間になんらかの人為的外乱が起こったとも考えられる。具体的には教育内容や英語学習動機における差などが挙げられ、これらの要因が学習の伸張の違いに結びついた可能性も考えられる。

6. 討論

6.1 能力レベルとスコア変動

上記5で報告した結果から以下のようなことが認められる。神戸大学入学から約8ヶ月後の国際文化学部1年生の英語能力は：

1. 入学時の英語能力が低い学生(TOEFL460以下)の場合は、8ヶ月間の間に英語能力が伸びた(平均で+20.39ポイント)。
2. 入学時の英語能力が中程度であった学生(TOEFL461以上480以下)の場合は、英語能力がわずかに伸びた(平均で+2.30ポイント)。
3. 入学時の英語能力が高かった学生(TOEFL481以上)は、英語能力が低下した(平均で-9.86ポイント)。

8ヶ月間の学生の英語学習に上記のようなスコアの変動があった原因としては、(1) テストの形式に慣れたことによる学習効果、(2) 平均への回帰(regression to the mean)、(3) 英語能力が実際に向上や減退したことによるスコア変動、の3つが考えられる¹²。観察されたスコアの変動が必ずしも英語能力の変動に由来するものではない可能性があるものの、その点を考慮しつつ得られたデータの解釈をしていきたい。

外国語・第二言語学習に影響を及ぼすような要因は、教室での授業、外国語学習適性、学習意欲、教室外での自習などのように多岐に渡るが、数々の要因のうちの多くを今回の分析では扱っていない。未知の要因が影響を及ぼしている可能性がある一方で、本調査に参加した学生に共通している要因の一つは8ヶ月間の英語授業である。受講8ヶ月後に通常予想されるのは、英語能力が大幅に向上しないまでも少なくとも減退しない状態である。ところが本研究に参加した者のうち、表6と表8に示したように、TOEFL1での上位層(なかでも541以上の高スコア取得者)は8ヶ月間の英語授業にも関わらず英語能力が減退しており、また中位層ではスコアが微増しただけである。一方で低位層は他の2つの能力層に比べ比較的能力が上昇している。共通英語の授業内容については調査を行っていないため明言はできないが、共通英語の授業は上位層の学習者にとっては平易すぎて能力の増強につながっ

ていないのではないかと推測される。

6. 2 セクション別能力変動

上記5. 2. 2で述べたように、 t 検定の結果、TOEFL1とTOEFL2のスコア間の変動のうち、リスニングと読解において有意な差があることがわかった。3セクション総合スコアと構造における差は有意とはならなかったが、リスニングについてはスコアが有意に上昇しており、読解においてはスコアが有意に低下している。

この結果については以下のような説明が考えられる。リスニングはおそらく高校の授業で十分に練習がなされていなかったため十分に伸びる余地が残されていた分野であったのだと推測される。そのため、大学入学後に受けたオーラルの授業で学生が練習の機会を得てリスニング力が伸びた可能性がある。その一方で、読解力がわずかではあるが低下していることについては、大学で受けた共通科目の英語購読の授業が、TOEFLのようなタイプの設問を解く読解力の増強には直接結びつかないためである可能性が考えられる。

表7中の百分位数は概算値ではあるが、それぞれのセクションのスコアの百分位数を見ると、構造や読解のスコアが30パーセントイルを越えているのに比べて、リスニングのスコアはTOEFL1で約12パーセントイルであり、スコアの上昇が多少見られたTOEFL2においても約18パーセントイルと、かなり低くなっている。このことから、(1)国際文化学部の新入生は3つのセクションの中でリスニングの力がかなり低く、従って(2)リスニングの力はまだ伸びる余地が大いにあり、リスニングのスコアを伸ばすことで、TOEFLスコアの上昇につなげられるだろう、と言える。

6. 3 英語能力向上への提言

神戸大学の共通授業科目の英語カリキュラムは近年改革され、新カリキュラムが平成14年度4月からスタートしたばかりである。その運用が軌道に乗って新カリキュラムの効果が表れるにはまだ数年かかるであろう。その間、学

生の英語能力の向上が比較的短期間であらわれるような方策がないものであろうか。今回の分析結果に基づいて考えられる方策を下に述べる。

(1) リスニング力の増強を計る。

国際文化学部新入生のリスニングのスコアは他のセクションのスコアに比べて低く、スコア向上の余地が十分にある（表7の百分位数参照）。また、高校ではこの分野の増強を目指した指導は集中的に行われていない可能性が高い。そのため、リスニング力を高めるような介入を行えば、他の2つの分野に比べて比較的短期間でスコアの向上が期待できるであろう。

(2) 能力別クラス編成によって授業レベルの最適化を図る。

高得点者のスコアが8ヶ月間の英語授業を受けた後に下落している問題については、上位層のレベルに対応できるように、現状よりも高度な内容を扱う授業が必要であろう。それを可能にするためには、上位層に限らず、レベル別にクラス編成を行う。レベル毎に適合した授業内容を提供することで、英語能力が増強されるだけでなく、学習内容への興味の衰退や学習意欲の減退を防ぐことにつながるであろう。

7. 結論および今後の展望

本稿で報告した内容は、国際文化学部1年生に2001年5月7日（参加者123名）と2002年1月11日（参加者73名）に二回実施したITP TOEFLのスコアのうち、両回共に参加した68名の結果を分析したものである。3つあるセクションのうち、全般的にリスニング力が劣っている傾向がみられた。一回目のスコアの平均が474.76で、二回目のスコアの平均が479.25であった。参加者を一回目 TOEFL のスコア別に460以下、461以上480以下、481以上の3群にわけ、一回目と二回目間にスコアがどれくらい上下したかを調査したところ、460以下の低位層は大きくスコアが向上し、また461以上480以下の中位層ではスコアの向上が見られた。しかしながら、480以上の上位層では、一回目よりも二回目のスコアの方が下落した。これらの事実から、比較的短いスパンで成果が現れるように国際文化学部新入生の英語能力を向上さ

せるためには、リスニング力の強化と高スコア者層が満足できるようなレベル別授業の提供の二つが対応策として考えられる。

最後に、本稿で発表した内容の解釈について注意を促したい。本稿で発表した内容は、あくまでも新入生約144名のうちの68名のスコアについて分析を加えたものであり、新入生全員の参加を得てデータを得た場合とでは結果が異なる可能性は否定できない。しかしながら68名という一部のデータに基づくとはいえ、本研究以前に国際文化学部新入生の英語能力を TOEFL のような標準テストを用いて計測したことがなかったことや、標準テストのデータに見られる傾向を分析し数値化した情報が存在しなかったことを考えると、本研究で報告した内容が持つ情報価値は低くはないと言える。今後は、今回の被験者に実施したアンケートの結果を分析して TOEFL データと併せて吟味していく。アンケートに含まれる学習動機などの調査結果を参考にすれば、英語授業以外にどのような要因が働いていたのかについて、より適切な結論を導き出すことができるであろう。また、その結論に基づき、より現実に即した英語能力向上への方策を提示することが出来るであろう。

-
- 1 1998年のTOEIC受験者数は公開テストと団体特別受験制度(Institutional Program)を含めて831,000人。2001年度まででTOEICを採用した企業や学校などの団体は約3,500以上(延べ総数)である(池野, 2002, pp.46-47)。TOEICはEducational Testing Serviceの下部団体であるChauncey Group Internationalの登録商標である。
 - 2 2000年7月から2001年6月の期間に、紙媒体で行われたTOEFL(TOEFL-PBT)受験者のうち日本語を母語とする者は34,645人であり、コンピュータを使ったTOEFL(TOEFL-CBT)受験者のうち日本語を母語とする者は60,114人であった。TOEFLはEducational Testing Serviceの登録商標である。
 - 3 実際に日本の様々な大学でそのような英語能力試験が採用されている。竹蓋ら(2002)は、TOEFLやTOEICを採用している大学で学会誌、試験実施団体機関誌、新聞などで大学名とスコアが公開された例を紹介している(pp.202-3)。また、

Johnson ら(1999)は文京女子大学 1 年生に SLEP テストを行った結果を『文京女子大学研究紀要』で報告している。

- 4 田中順子.(2002).「国際文化学部新入生の英語学習動機の分析と英語能力養成への展望」. プロジェクトメンバー：田中順子（研究代表者）、宇津木成介、沖原勝昭、水口志乃扶. このプロジェクトの本来の目的は、第一には国際文化学部に入学者の学生の英語学習について、入学時の英語力、英語学習の動機、伸ばしたい技能などについて調査をする教育的な目的であり、その第二は学生の英語能力の自己診断と実際の能力がどう関連し、それが時間と共にどう変化するか分析するという研究的なものであった。本稿は国際文化学部学内プロジェクトとして行った調査結果の中間報告書に再考察を加えたものである。
- 5 ETS は TOEFL の信頼性を “a slight modification of Cronbach's coefficient α ” を用いて予測している(Wainer & Lukhele, 1997, p.742)。この信頼性の算出方法では各設問が独立であることを前提とするが、TOEFL ではリスニングやリーディングセクションに見られるように、一つの問題文に複数の設問（小問）が設定されていることが多い。この場合各設問が独立であるという前提は成立し難い。Wainer & Lukhele (1997)は、相互依存した設問を持つ試験の信頼性を、1991年から1995年に使用された 4 種のテストフォームについて、Item Response Theory を用いて算出した。各設問が互いに独立だと仮定して算出した場合には、テスト全体の信頼性として 4 種のテストフォームすべてにおいて $\rho = .96$ が得られ、関連している設問をひとまとまりだと想定して算出した場合には、3 種のテストフォームで $\rho = .96$ が、1 種のテストフォームでは $\rho = .95$ が得られたと報告している。
- 6 Test of Written English は、現在では上記 3 セクションとともに、紙媒体の TOEFL-PBT でもコンピュータを使った TOEFL-CBT でも受験が必須となっている。
- 7 本稿では、テスト実施の形態別にスコアを区別する必要がある場合は TOEFL と ITP TOEFL を区別して表現する。そうでない場合には ITP TOEFL のスコアも通常の TOEFL のスコアも TOEFL スコアと表現する。
- 8 Educational Testing Service (2001)の Table 12 (p.11)から転載。
- 9 TOEFL-CBT でも同様の計算式を用いる(Educational Testing Service, 2001, p. 39).

- 10 学部入学を希望する者全員のスコアに基づくものである。受験者の母語は日本語に限らない。
- 11 Table 20. Paper-Based TOEFL Total and Section Score Means – All Examinees Classified by Native Language (Educational Testing Service, 2001, p.14)参照。
- 12 (1) は、実際には英語能力が向上していないものの、テスト形式に慣れたことによりテストスコアが向上することを指す。(2) は、より多くの群からデータをとるにつれ、それぞれの群の平均値がすべての群から得られる平均値にだんだん近づいていくことを指す。今回のようにデータを計測する機会が増えれば、1回目にスコア平均が低かった群は2回目の計測ではスコアが上昇し、1回目にスコア平均が高かった群はスコア平均が下がるという傾向がある。

(3) に基づくスコアの上昇、つまり純粋に英語能力が向上したことによるスコアの上昇分を計測するためには、(2) の要因はさておき、(1) の試験慣れによる学習効果を除いたスコア上昇分を検討する必要がある。(1) のような要因を排除するためには、最初のスコアを計測した1週間後に別のテストフォームを用いて再度スコアをとることにより、テスト形式への慣れによる学習効果をまず測定する。この最初のスコアと1週間後のテスト間のスコア差分を、研究期間終了時の最終テストのスコアと最初のスコアとの差分から差し引くことにより、テスト形式への慣れからくる学習効果を差し引いた純粋な英語能力の変動を測ることができる (Swinton, 1983 参照)。このような方策を本研究では採用していないため、TOEFL1 と TOEFL2 間で見られたスコア上昇分に英語能力の上昇以外の要因が含まれている可能性は否定できない。

参考文献

- 池野佐知子. (2002). 「TOEIC 喜びと苦しみの源泉」『AERA』2002年7月20日号, 46-51.
- 神戸大学国際文化学部自己評価委員会. (2001). 『神戸大学国際文化学部 現状と課題 (2001)』神戸大学国際文化学部.
- 竹蓋順子, 竹蓋幸生, 高橋秀夫, 土肥充. (2002). 「英語総合力養成のための CALL 教材の開発とその試用—科学研究費補助金による研究—」『ARELE』Vol.13,

- 199-208.
- Johnson, J. , アレン玉井光江, 加須屋裕子. (1999). 「SLEP テストによる英語能力測定 : 文京女子大学 1 年生の分析」『文京女子大学研究紀要』Vol.1, 141-162.
- Educational Testing Service. (1998). TOEFL: Computer-based TOEFL score user guide. Princeton, NJ: Author.
- Educational Testing Service. (2001). TOEFL: Information bulletin for computer-based testing. Princeton, NJ: Author.
- Educational Testing Service. (2001). TOEFL: Test and score data summary. Princeton, NJ: Author.
- Ewing, T. (2002, April). ETS launches three-phase introduction of next generation TOEFL. *News & Media*. Retrieved June 10, 2002, from <http://www.ets.org/news/04230201.html>
- Pierce, B. (1992). Demystifying the TOEFL reading test. *TESOL Quarterly*, 26, 664-691.
- Spolsky, B. (1990). The prehistory of TOEFL. *Language Testing*, 7, 98-118.
- Swinton, S. A. (1983). *A manual for assessing language growth in instructional settings* (TOEFL Research Report 14). Princeton, NJ: Educational Testing Service.
- Wainer, H., & Lukhele, R. (1997). How reliable are TOEFL scores? *Educational and Psychological Measurement*, 57, 741-758.

本研究は国際文化学部教育研究プロジェクト研究費の助成を受けて行われたものである。

本研究の実施にあたって、試験実施のための準備と監督のお手伝いをいただいた方々や草稿を読んで有益なコメントを下された方々に深く感謝申し上げます。