

TOEFL利用型英語カリキュラムに関する比較研究 —「RLGテスト」から浮かび上がる特徴と課題—

加 藤 千 博
(防衛大学校)

A Comparative Study on TOEFL-Based College English Curricula in Two Universities: Analysis with the Scores of Center Test, TOEFL, and “RLG Test”

Chihiro Kato
(National Defense Academy of Japan)

The aim of this study is to examine the problems of college students who follow a TOEFL-based curriculum. The curricula of two universities are compared using the scores of the Center Test (the National Center Test for University Admissions), RLG Test, and TOEFL. Although their respective English programs vary in many aspects, both schools require their students to take the TOEFL as their end-of-term exam for crediting or grading purpose. The participants in this study are 57 students from University A and 114 students from University B. The results show (1) that on the average, while the students of University A got lower scores than the students of University B in terms of the Center Test and RLG Test, they obtained higher TOEFL scores, and (2) that among the three skills (reading, listening, grammar) observed in this research, grammar has the greatest effect on the TOEFL scores of the students of both universities. These results suggest that the nature of the English curriculum in each school explains the difference in the test performance between the two sets of students covered in this study.

An “all-English class,” that is conducted using only English by both foreign and Japanese teachers, has on the one hand a positive effect on the students’ English practice and in turn on their TOEFL scores, and a negative effect on their grammar practice on the other. The combination of an “all-English class” and a “partial-L1-use class” according to students’ levels is worth considering. Setting achievement goals is one of the most effective methods to motivate students and to make them continue learning English. In addition, how to relate students’ self-study with a TOEFL-based curriculum is another key factor.

[キーワード：大学英語カリキュラム，センター試験，TOEFL，RLGテスト，到達目標，オールイングリッシュクラス，学習時間]

はじめに

大学における英語教育改革が叫ばれるなか，TOEFL

やTOEICなどの外部テストを英語カリキュラムに導入する大学が近年増加傾向にある。国際ビジネスコミュニケーション協会(2010a)によると全国で314校の四年制大学でTOEICが単位認定に用いられている。一方，国際教育交換協議会(2008)によると全国で141校の四年制大学でTOEFLが単位認定に用いられている。

齊田他(2009)は大学教育における外部試験活用事例を概観したうえでその目的を次の8点にまとめている。①国際的な水準による在学生英語力の把握、②習熟度別クラス編成、③英語授業の成績評価、④英語力の伸びの把握、⑤教育効果の検証、⑥卒業要件・進級要件としてのスコア認定、⑦単位の認定、⑧英語学習の動機づけ。これをさらに分類すると次の三つのグループに分けることができよう。Iスコアによって学生の成績、単位認定、進級が判定される：③⑥⑦、IIスコアによって学生のレベル、授業効果を測定する：①②④⑤、IIIスコアを学習上の励みとする：⑧。Iの分類は外部テストの英語カリキュラムへの積極的な導入であり、学生に対して英語学習を程度の差はあるが半ば強制するものと言える。

本稿では、分類Iのように外部テストを積極的に英語カリキュラムに導入している二つの大学の事例を比較し、それぞれのカリキュラムにおける利点と問題点、さらには両校の学生の英語学習上の課題点を探究する。本研究が対象とする大学はともに関東圏内の同県同市に位置する国公立大学法人であり、入学時の学生の英語レベルは非常に似通っている。そのため双方のカリキュラムの違いが入学後の学生の英語力の特性を色濃く反映している可能性が高い。ともにTOEFLを受験することがカリキュラムの一環となっているが、その利用方法と授業の進め方は大きく異なっている。本研究では、研究参加者の入学時の英語力を大学入試センター試験の得点から、授業による学習効果をTOEFLスコアから測定する。しかしながら、英語の知識を問うセンター試験と英語の運用力を測るTOEFLでは測定する英語力の質が異なってくる。そこで「RLGテスト」と呼ばれる英語の基礎力を測るテストを加えることによって、より客観的に両校の学生の英語力を比較できるようにする。

1. 研究の背景

国際ビジネスコミュニケーション協会(2009)の行った「就職活動をする学生に期待するレベルはTOEICスコアでいうとどのくらいですか」という質問に対して242校の大学から回答があり、「学部学科によって異なる」(23.6%)という回答を除くと、「600点～645点」が最も多く21.5%、次いで「500～545点」が20.2%であり、四年制大学での平均は569.4点($n=210$ 校)であった。それに対して、2009年度の四年制大学生のTOEIC平均スコアは公開テストで554点($n=217,310$ 名)、IP(団体受験)テストで441点($n=364,277$ 名)である(国際ビジネスコミュニケーション協会,2010b)。よって双方を合算した平均を算出すると483.2点となる。すると、大学側

が期待する平均スコア569.4点と実際の平均スコア483.2点の間には86.2点の開きがあることになる。しかし実際に就職活動を行うのは大学3,4年生であるので、2009年度IPテストにおける大学3,4年生の平均スコアを見てみると、3年生が484点($n=59,176$ 名)、4年生が506点($n=20,537$ 名)となっている。よって大学側の求める理想と現実にはなお3年生で85点以上、4年生で63点以上の開きがある。

このような理想と現実のギャップを埋める努力を各大学は果たして行っているのだろうか。同協会の「必修の英語授業時間は1年間に何時間ありますか」という質問に対して249校が回答し、最も多い回答が「90時間程度」で18.1%、次いで「45時間程度」が12.9%、平均では103.0時間であった。つまり、通年30週間、90分間の授業を週に2回行っている大学が最も多く、週に1回だけ行っている大学が次に多いことを示している。これは日本の大学での英語の授業数がいかに少ないかを示すデータであると言えよう。

この少ない授業数をTOEFLやTOEICなどの外部試験を導入することによって、齊田他があげる「⑧英語学習の動機づけ」を各大学は試みようとしているのかもしれない。学習意欲を喚起させることは英語教育において非常に重要であることに疑いはない。しかし動機づけの手段として利用する外部テストのための学習を単に学生自身による自律的な学習に委ねるだけでは教育機関として無責任と言えよう。これをいかに普段の英語授業と連動させていくかが各大学の抱える大きな課題である。そこでTOEFLをすでにカリキュラムに導入している二大学を比較し、そこから双方の学生及びカリキュラムの抱える課題を考察していく。

2. 研究の方法

(1) 参加者

公立大学法人A大学在学の1年生、2クラス57名と国立大学法人B大学在学の1年生、2クラス114名を対象とした。この二大学はともに関東地区にある国公立大学法人で、調査対象とした学生の英語力は非常に似通っている。ともに2008年4月入学の日本人大学1年生である。

対象学生の抽出にあたってはそれぞれの大学の平均的な学生となるように配慮した。A大学は前期、後期ともにTOEFL-ITPによるレベル別のクラス分けがされており、今研究の対象としたクラスは、1年生約930名中の順位がおおよそ460～520位に相当する2クラスである。よって1年生全体のほぼ中間レベルの集団である。一方

B大学はセンター試験の英語の得点によって学部ごとにクラス分けがされており、対象としたクラスは経営学部の4レベル中2番目と4番目のクラスである。経営学部の英語力は他学部比べて高いレベルにあるため、この2番目と4番目のクラスを抽出することにより学校全体の平均的な学生に近いデータを示すことができると思われる。

今研究において両校2クラスずつを対象としたが、クラスサイズが異なるため、57名と114名という異なる人数の母集団による比較となった。しかしながら後述するセンター試験、RLGテスト、TOEFLの結果において両グループの標準偏差は各テストの各項目において似通っており、*t*検定の結果においても有意差が示されており、A大学の57名というデータ数が必ずしも少なすぎて妥当性に欠くとは言えない。よってこれらのデータが両校の学生の特性を十分に反映するものとして本研究では議論を進めていくことにする。

(2) 調査時期

2008年度の後期(第2セメスター)に調査を実施したが、センター試験のスコアは入学前の2008年1月に受験したものであり、RLGテストとTOEFLは2009年1月に受験したものである。

(3) A大学とB大学の英語カリキュラム上の共通点と相違点

ともにTOEFLを利用しているA大学とB大学は他大学とは異なった独自の英語カリキュラムを実施している。以下の表1は両校の1年次の英語カリキュラム上の共通点と相違点を示している。

表1 英語カリキュラム上の共通点・相違点

	A大	B大
期末テスト	TOEFL-ITP	TOEFL-Level 2
クラス編成	習熟度別 (全学部横断的)	習熟度別 (学部単位)
教科書	統一テキスト (3レベル)	統一テキスト (1レベル)
TOEFLスコア	進級条件	成績評定
授業数	週3回	週2回
科目	L/R/G	L・R/S or W
クラスサイズ	25名	L・R: 60名 /S or W: 25名
指導言語	全て英語	日本人教員は日本語
TOEFL対策	行わない	対策授業(後期L・R)
授業の目的	実践的な英語力	受信能力/ 自己表現力/作文力

共通点は、A大学B大学ともに期末テストにTOEFL

を実施している点である。また両校とも習熟度別のクラス編成が行われており、教科書もほとんどが統一されている。

異なる点は、まずはTOEFLの位置づけである。A大学は明確な目標点を設定し、しかも目標点への到達を進級条件にしている。単位が認定されるためには500点以上のスコアが必要であり、この単位取得が3年次への進級要件となっている。他方B大学では成績評定(秀・優・良・可)に利用するのみで到達目標は示されていない。よってA大学の方が、基準点に達しないと留年する恐れがあるため学生へのプレッシャーは大きい。

二つ目には授業数の違いがある。A大学では週3回の授業があるが、B大学では週2回である。すると半期での学習時間はA大学が67.5時間であるのに対して、B大学では45時間となり、年間ではA大学が135時間でB大学が90時間となり45時間の差が生じる。先に示した国際コミュニケーション協会のデータによる大学での平均英語授業時間が年間103.0時間であったことから、授業時間に関してはA大学は多く確保していると言える。

三つ目の違いは科目の違いである。A大学は「リスニング」、「リーディング」、「グラマー」からなる週3回の授業を設定しているが、B大学は「リスニング・リーディング」と「スピーキング」もしくは「ライティング」からなる週2回の授業を設定している²⁾。原則的に「リスニング・リーディング」を日本人教員が担当し「スピーキング」と「ライティング」を外国人教員が担当している。

四つ目の相違はクラスサイズである。A大学は科目に関わらず25名前後の学生数に制限しているが、B大学では日本人教員担当の「リスニング・リーディング」を60名前後の大きなクラスサイズとし、その代わりに外国人教員担当の「スピーキング」と「ライティング」を25名前後に抑えている。

五つ目の違いは教室内での指導言語である。A大学では日本人教員外国人教員に関わらず授業は全て英語で行われ、学生も授業中は原則日本語を使用しないことになっている。一方B大学では外国人教員は英語で授業を行うが日本人教員は日本語で行っている。

六つ目の相違点はTOEFL対策の有無である。A大学では期末テストにTOEFL-ITP³⁾を利用するものの通常の授業ではTOEFLの題材を扱ってはいない。それに対しB大学では1年次後期の日本人教員担当の「リスニング・リーディング」クラスでTOEFLの問題演習を中心としたTOEFL対策の授業を行い、期末テストであるTOEFL-Level2⁴⁾の受験に向けて準備が施されている。

最後に両校の当該科目の授業の目的を比較すると、A

大学は実践的な英語力に焦点を絞り、リベラルアーツを学ぶために必要なレベルまでそれを引き上げることが目標となっている。B大学は「リスニング・リーディング」クラスでは、英語の受信能力に焦点を絞り、語彙力・速読力の向上をはかりながら、自然に話された英語が聞き取れるようになることを目標としている。また「スピーキング」は人前での自己表現力を、「ライティング」は正確な作文力を目的としている。

(4) 調査の手順

入学前の2008年1月に受けたセンター試験の得点が示す英語力が、1年間のそれぞれの大学での授業を経て、2009年1月に受けるRLGテストとTOEFLにどのように反映されるかが今調査での焦点である。

センター試験の得点は参加者に「筆記」「リスニング」および「合計点」を自己申告してもらった。推薦入試やAO入試によって入学した学生はセンター試験の得点を有していない場合がある。A大学では16名、B大学では15名がこれに相当した。これらの学生のデータはセンター試験の項目では除外した。

RLGテストはA大学B大学ともに後期（第2セメスター）の最後の授業（1月第4～5週目）で実施した。試験時間は60分間である。マークシート式の解答用紙を用いて行ったが、試験終了後すぐに学生自らが採点をし、その場で各自の弱点、課題がわかるように配慮した。授業終了時に得点を記した解答用紙のみを回収した。

TOEFLはA大学B大学ともに定期テストと位置づけられている。後期終了直後の1月末頃にA大学ではTOEFL-ITPが、B大学ではTOEFL-Level2が実施された。

センター試験、RLGテスト、TOEFLをそれぞれ大学別に平均値、標準偏差を求めた。平均値に有意な差がないかどうかを*t*検定を行い、大学間のレベル差を判定した。さらにRLGテストからTOEFLへの影響関係を探るため、相関係数を求めるとともに重回帰分析による検討を行った。

3. RLGテストとは

「鬼のA教授、仏のB教授」と称されるように大学における評価の恣意性が長い間批判されてきたが、それを是正する試みとして客観的な英語力評価に資するべく、大場（2009）が「RLGテスト（大学標準英語学力テスト）」と呼ばれる簡便な英語力測定テストを開発した。このテストは、「R（読んでわかる単語）テスト」、「L（聞いてわかる単語）テスト」、「G（文法の基礎知識）テスト」からなっており、Rテストは15分50問、Lテストは15分40

問、Gテストは30分60問のそれぞれ四者択一式⁵⁾で、トータル1時間を要し、各問1点で150満点となっている。

RテストはJACET8000の頻度順5000語をもとに、Lテストは4000語をもとに作成されており、Gテストは「大場シラバス⁶⁾」における「基本6品詞」「文」「動詞の拡充」「文の拡充」「動詞の転換」「文の転換」の6領域から各10問ずつ60題、初級、中級、上級レベルの問題が20問ずつになるよう配置されている（2010、大場）。

このRLGテストは、Reading語彙サイズ、Listening語彙サイズを測り、基礎的な文法項目の理解度を測定するテストであり、TOEICやTOEFLのような英語の運用力を測るテストではなく、英語学習者の基礎力、つまり潜在能力・基礎体力を測るテストと言える。しかしながらTOEFLやTOEICのようなテストとの高い相関があり（加藤他、2011）、そのうえTOEFLやTOEICのように事前の準備を必要とはせず気軽に受けることができる。TOEICは標準誤差を超えて得点が上下することが個々の学生にはしばしば見受けられるが（Childs, 1995；竹蓋他、2007）、RLGテストはいつ受けてもその時点での基礎力が正確に判定できる⁷⁾。TOEFLやTOEICのスコアレポートでは見えてこない特性、課題がこのRLGテストから浮かび上がってくる。よって本研究では、センター試験、TOEFLに加えて、第三の尺度としてこのRLGテストを使用することにする。

4. 結果

(1) センター試験の成績

A大学の参加者における大学入試センター試験、英語科目（筆記、リスニング）の合計得点は180～236点で平均は212.7点であった⁸⁾。一方、B大学の参加者における合計得点は176～238点で平均は219.1点であった⁹⁾。以下表2にセンター試験における筆記、リスニング、合計点の平均値、標準偏差、および*t*検定の結果を示す。リスニングの平均値においては両校の参加者間に有意な差は見られなかったが、筆記と合計の平均値はともに有意にB大学の方が高かった。よってセンター試験の得点においては若干B大学がA大学を上回っていることがわかる。

表2 センター試験

	A大 (n=41)		B大 (n=99)		<i>t</i> 検定結果 <i>t</i> 値
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
筆記	172.9	14.2	178.8	9.9	-2.28 A大<B大
リスニング	39.8	6.1	40.4	5.2	-0.58 <i>n.s.</i>
合計	212.7	15.6	219.1	11.1	-2.38 A大<B大

(有意水準5%とした両側検定)

(2) RLGテストの結果

Rテスト, Lテスト, GテストおよびRLG合計点に対する平均値, 標準偏差, *t*検定の結果は以下の表3に示すとおりである。A大学B大学の数値を見比べるといずれの項目においても非常に似通っている。*t*検定の結果では, Rテストにおいては有意差は見られなかったが, Lテスト, Gテスト, RLG合計においてはいずれも有意にB大学の方が高かった。よってRLGテストにおいてもB大学がA大学を上回っていることがわかる。

表3 RLGテスト

	A大 (n=57)		B大 (n=114)		<i>t</i> 検定結果	
	M	SD	M	SD	<i>t</i> 値	
Rテスト	37.9	4.0	39.0	3.7	-1.86	n.s.
Lテスト	25.9	2.9	27.5	3.4	-2.99	A大<B大
Gテスト	43.5	5.8	46.9	4.8	-4.05	A大<B大
RLG合計	107.3	9.3	113.4	8.6	-4.25	A大<B大

(有意水準5%とした両側検定)

(3) TOEFLの結果

A大学はTOEFL-ITPをB大学はTOEFL-Level2を期末テストとして用いている。異なったテストどうしを比較することには問題があるが, TOEFL-ITPとTOEFL-Level2はともにTOEFL-PBTの問題形式と得点基準に基づいており, スコアに互換性がある。よってここでは等質のスコアを算出しているものとして比較する¹⁰⁾。それぞれのTOEFLの平均値, 標準偏差, *t*検定の結果は以下の表4に示すとおりである。平均値ではA大学が16.3点の差でB大学を上回っており, 統計的にも有意な差であることが認められた。しかしB大学で用いられているTOEFL-Level2は500点満点であるためそれ以上の得点を測ることはできず, いわゆる天井効果のため, 実際の実力では平均値はもっと高い可能性がある。このB大学の参加者において500点満点の人数は8名であった。この8名はTOEFL-ITPもしくはTOEFL-PBTを受けていれば500点以上の得点を獲得していたことが予想される。そこで仮にこの8名の得点が520点であったとしてB大学の平均値を算出し直すと468.3点となり1.4点高くなる¹¹⁾。しかしそれでもA大学とは14.9点の差が残り, TOEFLスコアに関してはA大学がB大学を上回っている¹²⁾。

表4 TOEFL

	A大 (n=57)		B大 (n=114)		<i>t</i> 検定結果	
	M	SD	M	SD	<i>t</i> 値	
TOEFL	483.2	21.9	466.9	21.7	4.58	A大>B大

(有意水準5%とした両側検定)

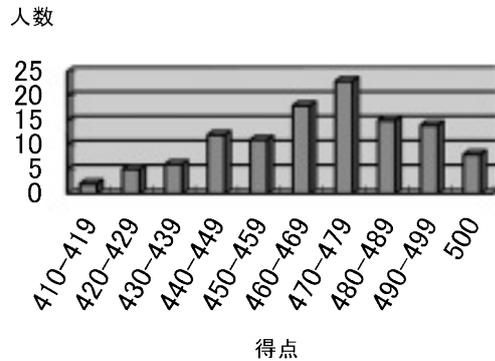


図1 B大 TOEFL-Level2得点群ごとの人数分布

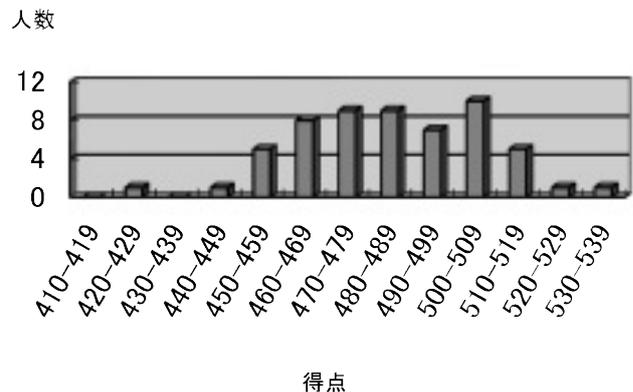


図2 A大 TOEFL-ITP得点群ごとの人数分布

上の図1, 2は, 両校のTOEFLの得点群ごとの人数分布を見るために10点ごとの得点群に分けたグラフである。図1のB大学では470点台を頂点としたいわゆるベルカーブを描いている。平均値が466.9点であることから460点台と470点台に人数分布が多いのはうなずける。一方, 図2のA大学では平均値が483.2点であることから480点台あたりを頂点としたベルカーブを描くことが予想されたが, 実際は500点台が最も人数分布が多くなっており, B大学のようにきれいなベルカーブを描いてはいない。

(4) RLGテストからTOEFLへの影響 (相関係数と重回帰分析)

Rテスト, Lテスト, Gテスト, およびRLG合計点とTOEFLとの相関係数は以下の表5に示すとおりである。ここでは各下位テストのなかではGテストがTOEFLとの相関が最も高いことがわかる。つまり, 英文

法の成績とTOEFLの成績の間に強い繋がりがあることになる。

表5 RLGテストとTOEFLとの相関係数

	<i>n</i>		Rテスト	Lテスト	Gテスト	RLG合計
A大	57	TOEFL-ITP	0.31*	0.27*	0.52**	0.55**
B大	114	TOEFL-Level2	0.32**	0.47**	0.50**	0.60**

*: $p < .05$, **: $p < .01$

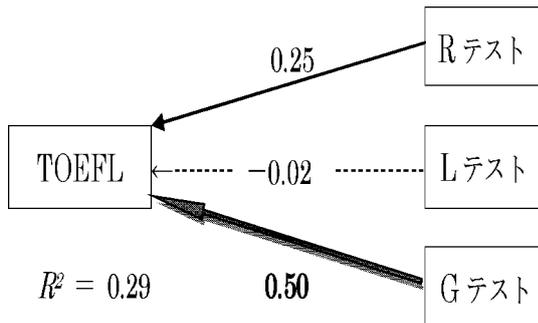


図3 A大学における重回帰分析結果

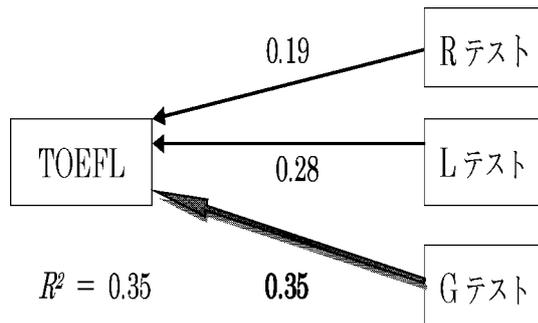


図4 B大学における重回帰分析結果

GテストとTOEFLとの影響関係をさらに探るため、TOEFL得点を従属変数、Rテスト、Lテスト、Gテストのそれぞれの得点を独立変数とした重回帰分析による検討を行った。上に示した図3はA大学における重回帰分析の結果である。調整済みの R^2 （説明率）は0.29であり説明力（回帰式のあてはまりのよさ）はやや弱い。回帰式を却下するほどではない。標準化係数に関しては、Lテストからの係数は-0.02とほとんど影響力はないと考えられるが、Gテストからは0.50となっており、TOEFLへの影響力が最も大きいことを示している。同様に図4からは、B大学においてもGテストからの標準化係数が0.35と最も高くなっていることがわかる。よって、A大学B大学の学生にとってはGテスト、すなわち文法において高得点を取ることがTOEFLでの好成績につながることを確認できる。これらの学習者集団においてはいかに文法学習が大切であるかが理解できよう。

5. 考察

以上の結果から、英語の知識（読みとリスニング）を測る大学入試センター試験と英語の基礎力（語彙と文法力）を測るRLGテストではB大学が上回っているのに対して、英語の運用力を測るTOEFLでは逆転してA大学が上回っていることが見てとれる。センター試験は2008年1月の結果であり、RLGテストとTOEFLはそのほぼ1年後の2009年1月の結果である。つまり入学前のセンター試験ではA大<B大であったのが、1年後のRLGテストでもA大<B大のままであり、英語力の差は変わっていないかのようなのであるが、TOEFLではA大>B大となっており逆転現象が起きている。先にも述べたとおり両校はともに関東地区の同県同市に所在する国公立大学であり、似たような特性の学生が在籍している。本研究の参加者においてはB大学の学生の方がセンター試験の得点が有意に若干高かったが、実際に学校全体のセンター試験の平均点も予備校等のデータではB大学の方が若干上回っており入学試験の難易度が高い。この研究の参加者はA大学B大学の中のごく一部の学生であるが、センター試験の得点から判断すると、各大学の平均を示す、代表的な学生であると言える。それぞれの大学を代表するような今研究の参加学生において、TOEFLスコアではセンター試験およびRLGテストとは逆転した結果が表れたことに関しては、それぞれの英語授業のカリキュラムが影響していることが大きな要因として考えられる¹³⁾。ここでは三つの要因、到達目標の有無と英語による授業の功罪、そして学習時間について考察することにする。

(1) 到達目標の有無

先の図2のTOEFL得点群ごとの人数分布で示したようにA大学では500点台が最も人数分布が多くなっており、B大学のようなきれいなベルカーブを描いていない。A大学は到達目標を500点と設定しているためその効果が、グラフに示されるように、500点台の人数分布の多さに表れているのではないだろうか。もし500点という到達目標がなければ、図2の500点台の得点人数が470点台、480点台に分散し、図1のB大学のようなベルカーブを描いた分布になっていたかもしれない。

図1と図2の得点群ごとの人数分布を比較すると二つの点が観察できる。一つは、450点以下の得点者の割合である。図1のB大学では410~449点の範囲に得点者が多く見られるが、図2のA大学では410~449点の範囲にはほとんど得点者が見られない。二つ目は、図2のA大学における500点台の得点者の突出である。500点が到達目

標であるためその得点をまさに超える地点の500～509点に得点分布が集中する現象が起きている。

するとA大学の到達目標500点という制度はある程度機能していると言えよう。それに対しB大学のTOEFLカリキュラムは450点以下の得点者の数を見る限り、それほど機能してはいないと言える。よってB大学にはA大学のように到達目標を設定するか450点以下の者を不可(単位不認定)とするなどの措置が必要かもしれない。A大学では3年への進級条件としてTOEFL 500点が設定されているが、2年への進級に関しては何も設定されていない。到達目標の設定による効果を考えれば2年への進級に際しても進級条件とはしないまでもTOEFLの到達目標を示すことは効果的と思われる。

(2) 英語による授業の功罪

A大学では英語授業を全て英語で行い、学生同士も教室内では英語を使うことを求められている。その効果がTOEFLスコアに表れていると考えられる。TOEFLのリスニングセクションの聞き取りやリーディングセクションのたくさんの文章を読みこなすには、単なる問題演習だけではなく、普段からの英語使用が要求される。授業を英語で全て行うことによってA大学の学生にはその対応がかなりできていることが想像できる。言い換えるなら、B大学の外国人教員による「スピーキング」「ライティング」授業と日本人教員による「TOEFL対策」授業がA大学と比較して効果が表れていないことになる。B大学の学生は基礎的な英語力はあるにも関わらず、TOEFLスコアにそれが反映されていない。しかし見方を変えれば、TOEFLでもっと高い得点を得られる素地は備えていることになる。

しかしその一方で文法の重要性も見逃せない。先のRLGテストとTOEFL得点の重回帰分析による検討の結果からはGテストがTOEFLスコアに対する影響力が最も大きいことがわかった。表3で確認したR・L・GそれぞれのテストにおけるA大学B大学間の平均値の差は、問題数の違いを考慮に入れても、Gテストにおいて最も開きが大きい。つまりA大学の学生はR・Lテストと比べてGテストの成績が最も良くない。これは文法力においてA大学の学生に何かしらの問題があるか、あるいはB大学の学生が非常に優れていることになる。B大学ではTOEFL対策の授業において文法セクションの問題演習を行っていることから文法問題への対応はしっかりとなされていると言える。一方、A大学では「グラマー」の授業も英語で行われており、学生にとっては文法に対する理解を妨げる一因となっている可能性がある。「現在完了進行形」を“present perfect progressive”と説明

を受けてもピンとこないのが学生の本音であろう。ましてや日本の中学、高校で学習する文法事項を理解していない外国人教員が「グラマー」の授業を担当すると、日本人大学生に対して適切な文法指導を行うのは難しい。ここに英語による英文法授業の限界が見えてくる。

A大学で見られるようにオールイングリッシュによる授業はプラスの面が多く、今後導入する大学が増加することが予測される。しかし英語を第二言語(ESL)としてではなく外国語(EFL)として学ぶ環境にある日本人大学生にとっては効果的に文法を習得することが英語力向上に資することは否定できない。よってコミュニケーション的なアプローチによる授業を展開しながらも部分的には母語である日本語を用いて効果的に説明を行うことも学習者のレベルに応じて考慮する必要があるだろう。Matsumoto(2011)はTOEFL-ITPで540点前後の上位クラスと470点前後の中位クラスにおいて、上位クラスではオールイングリッシュが、中位クラスではPartial L1 Use(部分的日本語使用)が効果的であると報告している。

(3) 学習時間

A大学の英語授業時間が年間135時間でB大学が90時間であることは先に論じたが、この時間の差がTOEFLスコアの逆転現象を引き起こしているとは結論づけるのは早計であろう。もちろん授業時間を多く確保することが学生の英語力向上に繋がることは否定されるものではない。どの大学でもできる限り英語の授業数を増やしたいのは同じであろう。しかしながら現実問題として容易に授業数を増やせるものではない。だからこそ各大学はTOEFLやTOEIC等の外部試験を導入して、少ない授業時間を学生自身による自発的学習に頼って補おうと試みているのである。よって重要なのは大学の設定する授業時間の他にどれくらいの時間をその学生が英語学習に費やしているかである。

A大学では単位取得条件として500点という明確な基準が設定されている。よって学生はその基準に達するために授業外で多くの自主学習を強いられている。それは単なるTOEFLの問題練習ではなく、eラーニングを使った発音練習やスピードリーディングの練習、さらにはGraded Readersを用いた多読の練習などが行われている。つまり授業外での自律的な学習に彼らは積極的に取り組んでいる。それに対してB大学では、同じくTOEFLを利用しながらも得点による目標数値設定がないため、学生の自発的な授業外学習は限られている。TOEFLの問題練習は十分行っていると言えるが、総合的な英語力を高めるような取り組みは十分にされている

とは言えない。この差がA大>B大というTOEFLスコアの要因の一つではないかと思われる。しかしRLGテスト、つまり英語の基礎力に関してはA大<B大のままであり、学習時間の差が基礎的な英語力を逆転させるまでには至っていない。つまりB大学の学生においては単に英語の運用力を養うような学習および練習が不足していると考えられる。

おわりに

これまで二大学の学生の英語力をセンター試験、RLGテスト、TOEFLのスコアを用いて比較し、それぞれの学生の課題と両校の英語プログラムの課題を考察してきた。A大学の学生は文法力に課題があり、オールイングリッシュでの授業の弊害が出ている可能性がある。レベルに応じて、オールイングリッシュのクラスとPartial L1 Use (部分的日本語使用) のクラスを設定することも考慮に値するだろう。しかし反面、英語で授業を受けるのに慣れているためTOEFLのような運用力を測るテストへの対応がなされている。またTOEFL 500点という明確な到達目標があるため500点台のスコアを獲得する者が多く、動機づけに役立っている。

ところがA大学では、TOEFL 500点を獲得しその学期に所定の出席数を満たすことによって単位が認定されるが、優・良・可といった教員による成績評定が行われない。そのため学生はTOEFL 500点を取ってしまえば、授業でのパフォーマンスは関係なくなり、ディモチベーション (意欲喪失) につながる可能性がある。実際A大学では学期途中でのドロップアウト (履修放棄) 者が少なくなく、英語学習にとって重要なモチベーションの維持と継続的学習が一部の学生の間で損なわれてしまっている。3年への進級条件としてのTOEFLの到達目標がうまく機能していることを考えると、1年次終了時点での到達目標を示すことで、このディモチベーションと継続的学習の問題は解決できるかもしれない。

一方B大学では実力があるにも関わらずそれがTOEFLのスコアに反映されていない。到達目標がないため積極的な動機づけがなされていないことがその一因と考えられる。到達目標の設定と成績評定に不可 (単位不認定) の可能性を含めることが検討に値しよう。ただしTOEFLのスコアを到達目標とするとそれが目的となってしまう、本来あるべきバランスのとれた四技能の習得が損なわれてしまう危険性がある。目標スコアはあくまでも英語力向上における一通過地点でなくてはならない。数値の設定とその位置づけは慎重に検討する必要がある。

B大学の学生はTOEFL対策の授業を行なっているぶん文法セクションには強いが、リスニング時の瞬時の聞き分け、反応、判断やリーディング時の短時間で多くの文章を読みこなす力は養われていないようである。TOEICやTOEFL-ITPにはスピーキングやライティングのテストは含まれていないが、リスニングとリーディングのテストからスピーキングとライティングの能力もある程度測ることができると言われている。つまりリスニングとリーディングのテストが間接的にスピーキングとライティングも含めた四技能を測定するのである (2010, 石川)。これは言い換えれば、TOEICやTOEFLのようなテストではリスニングとリーディングの練習さえ行っていればスコアが向上するというのではなく、スピーキングとライティングも含めた四技能全てを練習しておかないと思うようにスコアが伸びないことを意味している。B大学では「スピーキング」と「ライティング」クラスの一層の充実が目下の課題と言えるだろう。「スピーキング」と「ライティング」のクラスサイズを「リスニング・リーディング」よりも小さくしている割にはその効果が表れていないようである。同時に、少ない授業時間を補うために学生自身による自主学習を促すような課題の提供と、それをどうTOEFLを利用したカリキュラムと関連づけるかも大事な要素である。

今研究の対象とした二大学の学生は全国的な学生の平均レベルと比較すると高い英語力を備えた学生と言える。TOEFLのスコアを上げることが必ずしも総合的な英語力の向上に繋がるとは言えないが、英語力が向上すればおのずとTOEFLのスコアも上がることが期待される。今研究から両校の学生にとっては、しっかりとした文法力を備えることがTOEFLのスコアアップにつながる事が確認された。文法力を高めつつも、よりスピーキングとライティングを重視したアクティビティを行うことがこのレベルの学生には効果的であると言える。そしてこういった外部試験を導入することによっていかに授業外での英語学習時間を増やすことができるかが大きな鍵となる。語学学習は人生の一時期における一過性のものではなく生涯を通じて行われるべきものである。そのためにも学生時代に自律的な学習態度を身につけることが極めて重要である。

異なった大学間のプログラムを比較することは、そのプログラムの優劣をつけることが目的ではない。比較することでそれぞれの特徴と課題が浮かび上がり、改善へのきっかけとなる。大学の英語プログラムを改善することは一教員や一学科のレベルで行える容易なことではない。しかしながら各教員が問題点を認識しておくことで、

みずからの授業計画，クラス運営に生かすことができよう。大学においてよりよい英語の授業を行うためには，カリキュラム，学生の特性，英語教員としてのみずからの資質をしっかりと把握しておくことが必要となってくる。

注

- 1) A大学では習熟度に応じて3レベルのテキストが使用されている。B大学では推奨リストの中から各教員がテキストを選択することになっているが，ほとんどのクラスで同じテキストが使われている。
- 2) 前期に「スピーキング」を履修した場合は後期に「ライティング」を履修する。またはその反対である。
- 3) TOEFL-ITP (Institutional Testing Program) とはETSが提供する企業や学校などの団体向けのテストでTOEFL-PBT (Paper-based Test) の過去問が利用されている。
- 4) TOEFL-Level2はTOEFL-ITPより問題の難易度が低く500点満点で判定される。
- 5) 四択式の解答に際して偶然による正解の可能性を低くするために，第4選択肢には「(d)上のどれでもない」が設定されている。
- 6) Chomsky(1957)の助動詞群, Allerton, D.J.(1982)の転換子, Celce-Murcia & Larsen-Freshman (1999)のSMを参考に作成されている。大場昌也(2005)『学習英文法2005』
- 7) 加藤他(2011)はRLGテストおよびその作成方法に信頼性があるかどうかを証明するために，テストの二つのバージョンであるテストAとテストBの等価性を検証している。延べ約950名を対象に調査を実施し，R・L・Gそれぞれのテストについて，テストAとテストBの平均値，標準偏差，相関係数，*t*検定による検討が行われた。その結果，R・L・Gテスト全てにおける有意な正の強い相関から，R・L・GそれぞれにおけるテストAとテストBの等価性とテストの作成方法の信頼性が証明されている。さらに約230名を対象にRLGテストとTOEFL・TOEICとの相関を調査している。RLGテスト実施前後の1か月以内に受験された外部テストとのスコアが分析され，その相関係数が示されている。それによると，RLGテスト合計点との相関係数は，0.55から0.87までという有意な中程度から強い相関を示している。グループにより差はあるものの，RLGテストの合計点とTOEFLもしくはTOEICの間には相関があり，RLGテストの得点から各グループにおけるTOEFL・TOEICの得点予測がある程度可能であると結論づけている。
- 8) センター試験の点数は自己申告によるもので，記憶違いなどもあり完全なまでに正確なものとは言い切れない。推薦入試による入学者の中にはセンター試験のスコアを有していない者がおり，平均値は参加者57名中41名の平均である。
- 9) 同上。平均値は参加者114名中99名の平均である。
- 10) TOEFLを作成・運営するETSによると，TOEFL-ITPとTOEFL-Level2は500点以下の受験者にとってはほとんど同等のスコアを算出するとされている(ETS, 2000)。これを検証するため，Shaw(2005)は3グループ104名の大学生にTOEFL-ITPとTOEFL-Level2を1週間の間隔をおいて両方受験させている。結果，どのグループも後に受けたテストの方が平均点が若干高い傾向となった。これは「練習効果」による得点の向上であると考えられるが，テストの違いによる平均点の相違は確認されなかった。よってTOEFL-ITPとTOEFL-Level2は同等のスコアを算出するものと見なすことができる。しかしながらShawは500点以下のレベルの一人一人の学生に限って言えば，TOEFL-Level2の方がより正確なスコア判定ができると結論づけている。
- 11) 520点として算出した理由は，授業内で行った模擬TOEFLテストの結果，および習熟度のさらに高いクラスの学生のTOEFL-PBTのスコアを参考にした結果，当8名が520点以上の得点を獲得する可能性は極めて低いと判断したためである。
- 12) A大学では8割程の学生が，B大学では2割程の学生がTOEICのスコアも有しており，その平均点を比較した結果も同様にA大>B大となった。このことは「TOEFL-ITPとTOEFL-Level2が異なるバージョンであるため比較の方法に問題がある」という指摘に反論するデータとなる。しかしデータ数が少ないため本稿ではTOEFLのみを扱った。
- 13) 今研究において2008年度後期の授業に際してはA大学の「リスニング」「リーディング」「文法」とB大学の「リスニング・リーディング」の授業を同じ教員が担当しているため，教員の力量差による影響はさほどないと考えられる。

参考文献

- Childs, M. (1995) Good and bad uses of TOEIC by Japanese companies. In J.D. Brown & S.O. Yamashita (Eds), *Language Testing in Japan*, Tokyo: The Japan Association for Language Teaching.

- pp.66-75.
- ETS (Educational Testing Service). (2000) *TOEFL institutional testing program, 2000-2001, Japan Edition*, Princeton: Educational Testing Service.
- Matsumoto, Y. (2011) Influence of Partial L1 Use in the EFL classroom on English Learning - A Study Comparing the Intermediate and Advanced Levels. *JACET Summer Seminar Proceedings*, No.10, pp.41-45.
- Shaw, M. (2005) Is the Pre-TOEFL More Reliable Than the TOEFL? *Journal of Aomori Public College*, Vol.10 No.2, pp.55-61.
- 石川慎一郎(2010)「日本人英語学習者の目標言語能力特性：4技能の直接測定が示すこと」,『TOEIC Newsletter』No.108, pp.3-5.
- 大場昌也(2005)『学習英文法2005^T』よこはまティエロー出版局
- 大場昌也(2009)「大場昌也ホームページ」, 2009年4月1日検索 <http://www9.ocn.ne.jp/~bigarden/>
- 大場昌也(2010)「学士力(英語)測定のための「大学標準英語力テスト」の開発」,『第40回中部地区英語教育学会 学会設立40年記念石川大会 発表・提案要項』p.43.
- 加藤千博・田島祐規子・村上嘉代子・前川浩子(2011)「RLGテスト」の信頼性と妥当性の検討および形式的利用法に関する研究」,『中部地区英語教育学会紀要』第40号, pp.127-134.
- 国際教育交換協議会(CIEE)日本代表部TOEFL事業部(編)(2008)『TOEFL®テストスコア利用調査報告書2008年版』国際教育交換協議会(CIEE)日本代表部
- 国際ビジネスコミュニケーション協会(2009)『企業・学校における英語活用調査——2009年』財団法人国際ビジネスコミュニケーション協会TOEIC®運営委員会
- 国際ビジネスコミュニケーション協会(2010a)『TOEIC®テスト入学試験・単位認定における活用状況——大学院・大学・短期大学・高等専門学校——』財団法人国際ビジネスコミュニケーション協会TOEIC®運営委員会
- 国際ビジネスコミュニケーション協会(2010b)『TOEIC®テストDATA & ANALYSIS 2009』財団法人国際ビジネスコミュニケーション協会
- 斉田智里・小林邦彦・野口浩之(2009)「外部試験を活用した大学英語カリキュラム改革——茨城大学共通テストと外部試験との関連——」,『日本テスト学会誌』Vol.5 No.1, pp.95-105.
- 竹蓋幸生・与那覇信恵(2007)「英語教育カリキュラムの中の評価,その方法と問題点」,『文教学院大学外国語学部文教学院短期大学紀要』第7号, pp.143-158.

謝辞

本研究に際して,RLGテスト作成時の理論を横浜国立大学名誉教授の大場昌也先生(現金沢学院大教授)に直接ご教授いただき,統計分析に関しては金沢学院大の前川浩子先生にご助言をいただきました。ここに感謝を申し上げます。