

# 標準化テストと九州大学における英語教育

TOEFL ITP and English Education at Kyushu University

九州大学大学院言語文化研究院・教授 徳見 道夫  
立命館大学文学部・准教授 田中 省作

## Abstract:

Although people know that TOEFL ITP and TOEIC IP are standardized tests of English, they do not understand precisely the meaning of “standardized tests” nor the differences between them. This paper shows the differences clearly in order to furnish a basis for discussion of the TOEFL ITP scores of students at Kyushu University. Drawing on the analysis, the paper then proposes suitable methods for teaching English to students at the University.

キーワード：英語標準化テスト，TOEFL，九州大学における英語教育

Keywords: English Standardized Test, TOEFL, English Education at Kyushu University

## 英語標準化テスト(1)

英語標準化テストの代表的なものは、TOEFL (Test of English as a Foreign Language) と TOEIC (Test of English for International Communication) である。標準化テストとは、偏差値を用いてスコアを解釈しやすい数値に変換操作するテストのことを意味する。TOEFL ITP を例にとって説明すると、raw score (素点 = 正答数) を換算表で scaled score に換算したものが、各セクション (Listening Comprehension, Structure/Written Expression, Vocabulary/Reading Comprehension) のスコアとなり、さらに各セクションの合計を10倍して3で割ったものが total score となる。<sup>1)</sup> 換算表は

\* 本論文の執筆においては、3年生以上を対象とする「学士課程における英語教育の改善に関する調査 (TOEFL ITP)」のデータを使用した。その調査を実施され、データをご提供いただいた丸野俊一理事 (教育担当) には厚く御礼を申し上げます。

<sup>1)</sup> 高等教育開発推進センターの林篤裕教授から、標準化テストについて、統計学専門の立場から、次のように丁寧な説明をいただいた。

英語標準化テストでいう「標準化」は、統計分野でいう平均や標準偏差を揃える変換を指すのではなく、item (設問) ごとのスペックを事前 (pre-test) に把握しておき、新たに得られた受験者の成績がどこに位置しているかを示す指標を作るための技術を指す。ここでいうスペックには、難度 (difficulty, 易しい/難しい) と識別力 (discrimination, 識別できる/できない) の2つの指標を用いることが多い。TOEFL や TOEIC は受験者ごとに (もしくは試験会場ごとに) item が異なって出題されるが、偶然易しい item ばかりが出題された受験者が、それらに正解して加点したとしても、そのまま高いスコアを修めるわけではなく、既に調査が完了している過去の item ごとのスペックと比較して、受験生全体の中の位置を示すための技術とでもいえば、理解し易いかもしれない。個々人で異なる item が出題されるために、単なる正答数では成績の高低を比較できないのは、このような仕組みが内在しているからである。

なお、日本国内で行われている実用英語技能検定 (いわゆる英検) は、各級の実力範囲に幅があり、細かで正確な実力判定とはなりにくいこともあるが、標準化テストの一つである。中学生や高校生が各レベルの試験を受験し合格することは、英語を学ぶインセンティブになるので、英検の存在は日本の英語教育にとって有益である。また国連英検も特 A 級から E 級までの6段階のレベルが設定された標準化テストの一つであると考えられる。因みに、山下邦明言語文化研究院長 (2010年度および2011年度) は、国連英検の創始者の一人であり、「国連英検名誉特 A 級」を保持している。

問題によって値が変わり、問題の難易度によって得点が左右されないように操作される。従って、本稿で実施しているように、ある時期に受けた TOEFL 試験のスコアと別の機会に受験した試験結果を比較・検討することができる。

TOEFL と TOEIC の目的の違いを簡単に述べると、TOEFL は北米大学・大学院に留学する際、大学で英語を使って学習や研究が可能かどうかを計測するための試験であり、TOEIC は企業で働くために必要な英語力を測定する試験である。そこで北米の大学・大学院に留学するには TOEFL の試験を受験するし、企業に就職するには TOEIC の試験を受験することになる。九州大学で毎年実施されている英語標準化テストは、TOEFL ITP であるが、ITP とは Institutional Testing Program の略で、企業や大学が団体に受験する際に用いるテストである。TOEFL ITP の試験問題は、TOEFL の本試験で使用された過去の問題を使用している。

TOEIC と TOEFL の決定的な違いは、筆者らは読解問題にあると考えている。TOEIC の読解問題となる文書は、ビジネス界で使用されるメールや広告あるいは説明文などが主で、内容は簡潔で全体の文章も短いものである。一方、TOEFL の読解問題はアカデミックな内容に重点が置かれ、北米の大学で行われる講義に類するものが多い。例えば、標準的な TOEIC の読解問題は、以下のようなものである。

Hollywood is no longer just in California. Today the entertainment industry is finding new homes in Europe, Latin America, and Asia. The American media and communications industries are looking all over the globe for new opportunities. Although many companies are investing in the fast-growing European media industry, many industry executives believe the biggest long-term opportunity is in China and other countries in Southeast Asia. The potential market is huge over 310 million people in the European community, but over 650 million in the Pacific Rim.<sup>2)</sup>

全体の文章は短く、内容的にも分かりやすい英文であり、日常的に使用される英単語の知識があれば、十分に対応できるものである。その他に、新聞、雑誌の英語文章を問題として出す場合もある。因みに、TOEIC テストで基準となる成績は730点で、この点に達すると英語による日常的意思疎通はできると考えられている。<sup>3)</sup>

## 英語標準化テスト(2)

一方、TOEFL の読解問題はアカデミックな内容が多い。これは北米の大学生活では、英語テキ

<sup>2)</sup> Lin Lougheed, *Barron's TOEIC: Test of English for International Communication* (Barron's, 2007), p. 316.

<sup>3)</sup> 筆者らは標準化テストのスコアが英語コミュニケーション能力を正確に反映しているとは考えていない。この点については、菅原克也氏も『英語と日本語のあいだ』(講談社現代新書, 2011年)の中で、「現在、英語の能力試験 TOEIC, TOEFL といった試験の点数が一人歩きする状況がある。「コミュニケーション英語」の能力とは、英語の能力試験で高得点をあげる能力だという、逆立ちした理解すらみられる。英語によるコミュニケーション能力をじゅうぶんに備えた人は、試験でも高得点をあげるであろう。だが、試験で高得点だったからと言って、実際のコミュニケーションの場で期待される能力が発揮できるとはかぎらない。」(pp. 5 6) と論じている。まさに正鵠を射た指摘である。

ストを正確に早く理解する能力が重要となるからである。次に標準的な TOEFL の読解問題を検討してみよう。

The words “culture” and “society” are used interchangeably by many people, but their precise meanings are a little different. Society can be defined as the organized interaction among people living within a geographical or political boundary that is determined by culture. Culture, meanwhile, is the shared way of life of a particular group of people. It is composed of the values, beliefs, behaviors, and physical objects that the group shares.

Sociologists regard culture as possessing two distinct aspects: nonmaterial and material. The nonmaterial aspect is all of the intangible creations of a people. These are the culture’s values, ideas, and beliefs. They can be anything from egalitarianism to Taoism. The material aspect, meanwhile, is all of the tangible things that a culture creates — anything from pyramids to lawn mowers. Looking more closely, we begin to see that culture is, in fact, a complete social heritage.

Since ancient times, human cultures have evolved in unique and astonishing ways. For many thousands of years this happened very slowly, but in recent times, it has been happening more quickly. A graph depicting the timeline of various world civilizations could show this clearly. Many of the recent changes are not a result of adaptation to environment, like the kind animals make. (Animals do not create culture.) These changes are the original and unique creations of particular groups.

( 1 パラグラフ 7 行省略 )

Culture programs not only our pace, slow or fast, but also our reasons for choosing the pace. Cultures which choose to evolve at the pace of the natural environment do so because they choose to conserve resources. Cultures which choose speed also choose unlimited use of resources. No culture is intrinsically “natural,” but these two choices made by cultures in using resources produce groups of people whose lifestyles are directly at odds. One result of this is that we frequently hear of travelers today experiencing extreme culture shock.

A modern person finding themselves in an environmentally-based culture may need, for example, to learn to accept nudity as a part of daily life or eat foods that are shockingly cathartic to a body accustomed to eating processed foods. A traditional person finding themselves in a modern culture will likely be overwhelmed at the dizzying pace of modern life and feel bereft of the supportive matrix of a close-knit community of people, plants, and animals. In a modern culture, we are not likely to encounter behavioral demands that have environmental determinants. The required change has nothing to do with adjusting to the climate but has primarily to do with the need to find approval and a sense of belonging to the social group.

( 2 パラグラフ 11 行省略 )

The need exists for modern humans to align culture to the natural environment. The key to this change lies in re-examining and re-formatting its nonmaterial aspects — its values, beliefs, and behaviors. It also involves managing the material aspects differently.<sup>4)</sup>

---

<sup>4)</sup> Moraig Macgillivray, *et al.*, *Mastering Skills for the TOEFL iBT: Advanced* (Compass Publishing, 2006), p. 150.

長々と英文を引用した理由は、実際に TOEIC や TOEFL の試験問題を見たこともない人々が、英語の標準化テストについて滔々と論じていることがあるからである。実りある議論をするためには、まずテストの内容を正確に理解する必要があるので、あえて長文を本稿に引用した。

上記の TOEFL 英文を前掲の TOEIC 読解問題と比較すると、その分量と内容に関しては、全体の文章はかなり長いし、内容も抽象的かつ学問的である。この英文を正確に早く理解するためには、<sup>5)</sup> 英語の実力はもちろん必要であるが、日頃からの読書（英語と日本語）と思索が必要である。上記の文章に関する問題は、内容理解を確かめる問題が中心であるが、最後の問題は英文全体を理解していないと解答できないように工夫されている。<sup>6)</sup>

### 英語標準化テスト(3)

TOEIC と TOEFL の違いは、リスニング問題でも顕著にあらわれる。TOEIC の全体の英文は比較的短く、ビジネスの内容が大半であるが、TOEFL では文章の分量が多く、内容的にも難しい。日本人はリスニング英語に関しては集中力が長く持たないので、TOEFL のリスニング問題は苦手なようである。<sup>7)</sup>

標準的な TOEIC のリスニング問題のスク립トを見てみよう。

Although science can do many things, it can't change the weather. So the only defense people have against the fury of nature's winds is the weatherperson's warnings. To get a clear picture of our cloudy atmosphere, meteorologists have plenty of tools. They launch satellites into space to give us a God's eye view of our weather patterns. They target tornadoes with radar to determine its course. They even fly into hurricanes and seed the turbulent clouds with sensitive instrument packages.<sup>8)</sup>

問題文は簡潔な英文であり、全体の分量もそれほど多くないので、理解しやすいと考えられる。内容的にも、日本語で天気予報を聞いておれば、推測はそれほど困難ではない。ただ動詞 seed という単語は、日本人の英語学習者には馴染みがない言葉であろう。

次に標準的な TOEFL のリスニング・スク립トを見てみよう。

<sup>5)</sup> 本文で「正確に早く理解するためには」と書いたが、時間をかけて英文を正確に読み取るとは比較的簡単であるが、TOEFL テストでは読み取る「早さ」も要求される。読解問題が最後まで解けなかったという TOEFL 受験生が多いのは、この読解能力の「早さ」がないためである。

<sup>6)</sup> とところで、英文法に詳しい読者は気づいていると思うが、第6パラグラフにある最初の文 A modern person finding themselves in an environmentally-based culture may need...は、単数の A modern person を複数の themselves が受けている。これは gender への配慮のためである。himself と書けば男性中心の世界に逆戻りになるし、himself/herself ではごこちない文章となる。

<sup>7)</sup> 大学入試センター試験に英語リスニング問題が導入されて以来、リスニングに力を入れる高等学校が増えたせいか、TOEFL のリスニング問題にそれほど戸惑う学生はいないようである。これは筆者らの主観的な観測なので、客観的な指標がほしいところであるが、幸い、センター試験のリスニング問題導入と九大生の TOEFL スコアとの関連を、高等教育開発推進センターの林篤裕教授と調査中である。近くその調査結果を発表する予定である。

<sup>8)</sup> Wayne I. Phillips, 大谷啓明, 『TOEIC TEST 完全マスター600』(メディアファイブ, 2003年), p. 192.



We've seen that polysemic words are words that have numerous meanings. Well, there is a really special group of polysemic words, and this special group consists of words that have not just different meanings but opposite meanings. Think about this, that there can be one word in English that has not just different meanings but meanings that are, in some sense, opposite.

Think about the words “bolt” and “fast.” If you bolt something, perhaps you bolt your door, then this means that it's locked fast, or fastened and cannot move or open. However, if someone decides to bolt, then he or she is running away, and running away very fast. So the word “bolt” can mean “locked and not moving” or conversely can mean “moving very quickly.” The same can be said of the word “fast.” This word can also mean “locked and not moving” or “moving very quickly.” Another interesting word of this type is the word “sanction.” Just think about what this word means. If you sanction something, it means that you permit it. However, if you put a sanction on something, then it means the opposite. In this case it means that you don't permit it.

OK, keep in mind that these words that you just learned about, “bolt” and “fast” and “sanction,” are polysemic words because each can have different meanings. But they are special kinds of polysemic words because their meanings aren't simply different; their meanings are opposite in some sense.<sup>9)</sup>

読解問題と同じように、リスニング問題でも TOEIC よりも TOEFL のほうが長い文章が多い。そのため注意を集中して聞く必要がある。また polysemic という単語を知っている英語学習者はあまり多くないので、この意味を内容から類推しなければならない。polysemic という単語は、日本で最大の英和辞書である研究社『新英和大辞典』第六版にも収録されていない。そこには名詞の polysemy が載っているだけである。Oxford 大学出版の *Shorter Oxford English Dictionary* (CD-ROM 版)では、“of or pertaining to polysemy; having several meanings, exhibiting polysemy” と解説されているが、かなり特殊な単語であることは間違いない。

以上、TOEFL と TOEIC の違いを比較・検討してきた。結論として言えることは、TOEFL の問題を中心に学習していけば、TOEIC の問題は解答できるが、その反対は困難であるということである。換言すれば、TOEIC 問題を解く訓練をしておいても、TOEFL 試験対策にはあまり役に立たないことになる。どちらの試験にも対応できるように、九大生はまず TOEFL 問題に習熟しておくべきである。

ただ強調しておきたいことは、九州大学における英語教育は、標準化テストのためだけの教育ではないということである。「九大憲章」に謳われているように、「九州大学の教育は、日本の様々な分野において指導的な役割を果たし、アジアをはじめ広く全世界で活躍する人材を輩出し、日本および世界の発展に貢献することを目的とする」(九州大学教育憲章第 2 条)のであるから、英語教育もその理念達成に貢献することは当然である。九州大学における英語教育の目的は人材の養成であり、決して標準化テストの成績向上だけを目指すものではない。

なお、TOEFL ITP にはライティングとスピーキングのテストがない。リスニング、文法、読解

---

<sup>9)</sup> Deborah Phillips, *Longman Preparation Course for the TOEFL TEST* (Pearson Longman, 2006), p. 584.

に関する4択問題だけでは(選択問題の成績は「偶然」も関係してくる)、英語コミュニケーション能力は正確に計測できないので、九大学生には本試験である TOEFL iBT (Internet-Based Testing) を受験することを勧めたい。実は、数年前から九州大学で TOEFL iBT が実施できるように、関係者と協議・調整をしているが、様々な障害のためにまだ実施できないでいる。九州大学で TOEFL iBT の試験が実施できれば、多くの九大学生に益することになるであろう。

#### 九大学生の入学時と入学後の TOEFL スコア

九州大学の TOEFL 成績分析については、2007年に志水俊広准教授が、『大学教育』第13号に「英語標準化テストに見る九大1年生の英語力」というテーマで書いているし、本論文の筆者の一人である徳見も、2010年の『言語科学』第45号(九州大学大学院言語文化研究院言語研究会)に「標準化テストに見られる九大学生の英語能力の変化」を上梓している。さらに徳見は、2011年3月刊行の『言語科学』第46号でも、九大学生の TOEFL の成績を分析している。<sup>10)</sup>

上記3編の論文はいずれも入学時と1年後期の TOEFL スコアを分析・検討したものであるが、本稿では2009年度と2010年度に実施した3年生以上の九大学生が受験する TOEFL のスコア(以降、単純に入学後の TOEFL スコアと記す)と同学生が入学時に受験したスコアを比較・検討する。3年生以上が受験する TOEFL は、両年度とも希望者に対するもので、九大学生全体を母集団とした標本と考えるには無理があることには留意されたい。ただ、このように TOEFL を希望して受験するという態度は、九大学生の中では比較的英語に対して積極的な層であることを示している。また、本稿では2009年度と2010年度の入学後の TOEFL スコアを年度ごとに分析することとした。両年度のデータをマージして、より大きな一つのデータとして分析することも意義のあることではある。<sup>11)</sup> その一方で、年度によっては教学上の相違だけではなく、学生層や英語学習に対する態度等に多少の揺れが存在することも事実である。ここでは、各年度をそれぞれに分析し、両年度で共通する傾向を見出すことで、九大学生のより普遍的な傾向を表出することにした。なお、以降で示す数値データの中で、平均や割合(比率)については、適宜、統計的仮説検定を適用しており、その有意確率が0.05未満の場合を「有意である」と表現している。<sup>12)</sup>

まず、TOEFL を受験した3年生以上の学生数であるが、2009年度は234名、2010年度は374名である。ただし、この数には入学時あるいは入学後の TOEFL の一方のみを受験した学生も含まれている。本稿では、受験した学生全員ではなく、入学時・入学後の2回の TOEFL を受験した学生のみを分析対象としている。該当する学生は、2009年度では201名、2010年度では127名である。

TOEFL とその下位項目のスコアについて、年度別に入学時・入学後・その変化(入学後のスコアから入学時のスコアを減じた値)の平均・SDを表1に示す。なお、変化の平均については、その値が正であった場合に「スコア変化の母平均が0よりも大きい」を対立仮説とした対応ありのt検定(片側検定)を、負であった場合に「スコア変化の母平均が0よりも小さい」を対立仮説と

<sup>10)</sup> 徳見道夫・田中省作「英文法コーパス構築の有用性」『言語科学』第46号(九州大学大学院言語文化研究院言語研究会, 2011年), pp. 61-74を参照のこと。

<sup>11)</sup> 2009年度と2010年度をマージしたデータ分析は、別稿にて報告する予定である。

<sup>12)</sup> 検定の有意水準を0.05にした、という意味である。

した対応ありの t 検定 (片側検定) を適用した。なお、以降 LC は Section 1 (Listening Comprehension), SWE は Section 2 (Structure and Written Expression), VRC は Section 3 (Vocabulary and Reading Comprehension) である。数値に付した\*\*\*は  $p < 0.001$ , \*\*は  $p < 0.01$ , \*は  $p < 0.05$  で、無印は n.s. を意味する。

表 1 : TOEFL のスコア

	2009年度			2010年度		
	入学時	入学後	変 化	入学時	入学後	変 化
Total	455.9 33.9	453.2 39.2	- 2.7 34.9	464.7 32.8	455.6 45.8	- 9.2** 35.7
1 . LC	45.00 3.70	44.74 4.30	- 0.25 3.77	45.50 3.54	45.57 4.44	0.07 3.73
2 . SWE	45.43 4.08	44.37 5.29	- 1.06** 5.44	46.54 4.37	45.35 5.70	- 1.18* 5.16
3 . VRC	46.36 5.01	46.80 4.81	0.44 5.48	47.42 4.26	45.76 6.20	- 1.66** 6.23

図 1 と図 2 にそれぞれ2009年度と2010年度における各学生の Total スコアを階級幅10点とした際の相対度数を示す。両年度とも入学時には多少の非対称性は認められるものの、平均を中心に概ね対称的に分布をしている。留学に必要とされる500点以上を得ている学生は非常に少ない。2009年度については、入学時3割程度を占めている410から430点の度数が入学後には低くなり、入学時には極めて少なかった390~400点と460点に増加が認められ、対称性が崩れ分化の傾向が見られる。2010年度については、逆に入学時に450点を境に分化している傾向が、入学後にスコアが分散し、対称的な分布の傾向を形成している。

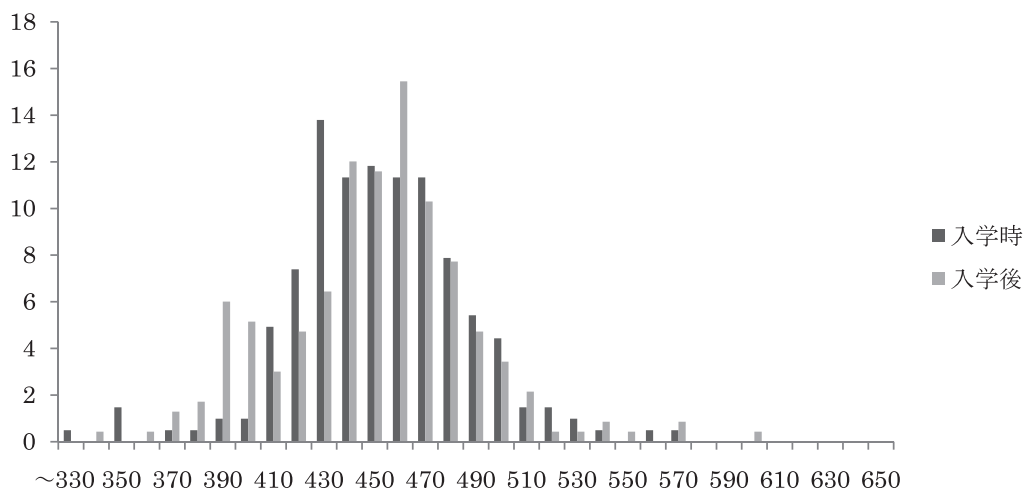


図 1 : 2009年度 Total スコアの分布 (縦軸は割合[%]・横軸は Total スコア)  
横軸のラベル  $\alpha$  は、入学時の Total スコアが「 $\alpha$  より大きく、 $\alpha + 10$ 以下」を意味する。

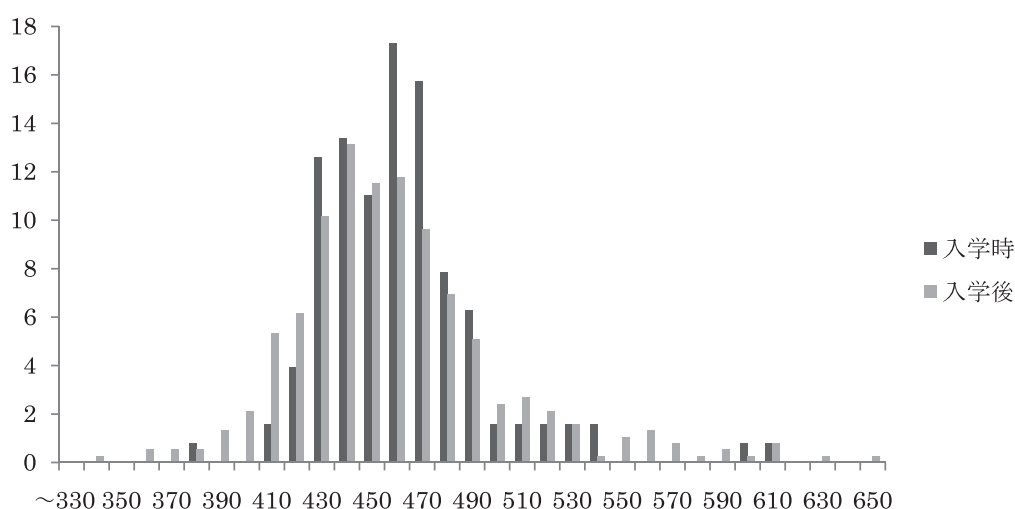


図2：2010年度 Total スコアの分布（縦軸は割合(%)・横軸は Total スコア)

入学時のスコアと入学後のスコアの変化を見なおしてみよう。まず Total スコアについては、いずれの年度も入学時に比べて入学後はその平均を下げており、2009年度は有意ではないものの2.7点、2010年度は9.1点も下げている。TOEFL 成績という観点のみから見れば、平均的には入学時から英語の力は大きく変化しない、もしくはやや下げていることが分かる。下位項目については、統計的に有意であった箇所注目すると、2009年度については SWE が1.1点、2010年度については SWE が1.2点、VRC は1.7点下げている。下位項目の平均1点の上下動は、一見すると小さなものに思えるが、あわせて SD にも着目すると、かなりの散らばりが見られ、大きく点数を落としている割合が高いことが分かる。

SWE については、両年度で共通して有意にスコアを下げることから、九大生<sup>13)</sup>の入学後の英語に対する関心や学習項目については、SWE に関わる内容から優先的に精査してみる価値はあるだろう。また、VRC については、2009年度ではほぼ変化がなかったのに対して、2010年度では大きくスコアが下がっている。この点については、次々節において入学時の Total スコアごとに大まかに区分して改めて取り上げる。

### 入学時の Total スコアから見た変化

TOEFL を受験した九大生について、両年度とも Total スコアの平均という点では入学時から有意な向上は見られず、2010年度についてはむしろ下げる結果となっている。このような現象をより正確に捉えるべく、ここでは入学時の TOEFL スコアを粗く区分し、スコア変化を精査してみる。<sup>14)</sup>

まず、図3に2009年度の入学時 Total スコアと入学後のスコア変化の散布図を示す。直感的には、入学時に400点に満たない学生は概ね入学後のスコアが上がる一方で、入学時のスコアが高くなるにつれて、入学後はスコアを下げる学生が多くなっていることが分かる。実際、入学時のスコアと

<sup>13)</sup> 3年生以上の TOEFL 受験の経緯から考えると、これを九大生全体の傾向と捉えることはできない。

<sup>14)</sup> ここで示した分析以外にも、受験者のプロフィール等も考慮したさまざまな分析が考え得る。今後、その他の分析についても報告する予定である。



入学後のスコアとの変化の相関をとってみると、2009年度では  $-0.35$  ( $p<.001$ ) で弱い正の相関関係が認められる。この傾向は、入学時の Total スコアを分割してみると、よりはっきりとしてくる。

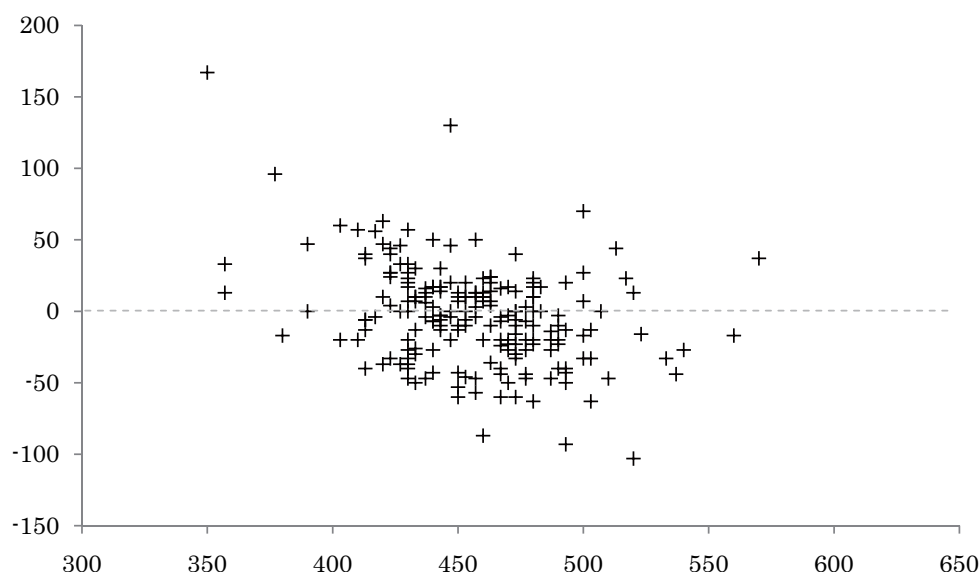


図3：入学時の Total スコア[横軸] × 入学後の Total スコアの変化[縦軸] (2009年度)

入学時の Total スコアを恣意的ではあるものの、400点以下と500点より高い層、400点から500点の間は20点刻みでスコアの区間を構成し、各区間に対するスコア変化の平均とSD、学生の数 (n)、スコアが上がる (変化が0より大となる) 学生の数 (m) とその割合 (m/n) を算出した。スコア変化の平均には、表1の変化に対する t 検定 (片側検定) を同様に適用した。スコアが上がる学生の割合については、「母比率が0.5と異なる」を対立仮説とした母比率の検定 (両側検定) を適用した。その結果を表2に示す。なお、区間 (a, b]は「aより大きくb以下」であることを意味する。その表2によれば、入学時のスコアが (0, 400]である学生のスコア変化の平均は有意に0よりも大きい、つまりスコアが有意に上がり、(460, 480], (480, 500], (500, 550] ではスコア変化が有意に0よりも低く、散布図から得られる直感が、全区間ではないものの統計的に確認される。

表2：入学時の Total スコアと変化 (2009年度)

入学時のスコア	平均	SD	n	m	m/n
( 0, 400]	48.4*	63.8	7	5	0.71
(400, 420]	14.0	36.6	16	8	0.5
(420, 440]	1.9	29.0	47	27	0.57
(440, 460]	- 0.4	33.9	45	23	0.51
(460, 480]	- 11.1***	25.0	49	16	0.33*
(480, 500]	- 16.0*	33.1	22	5	0.23*
(500, 550]	- 23.0*	38.5	13	3	0.23
(550, 677]	10.0	38.1	2	1	0.5

同様に、図4に2010年度の入学時の Total スコアとその変化の散布図を示す。2009年度と比べると、ランダム性が強くなり、実際、相関係数を計算してみると - 0.11 (n. s.) で、ほとんど相関が認められない。

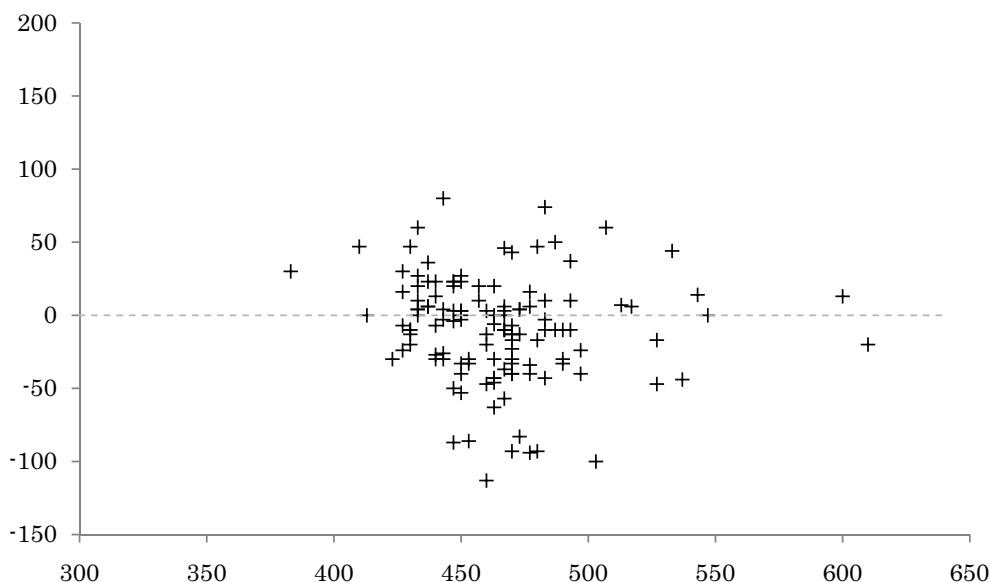


図4：入学時の Total スコア[横軸] × 入学後の Total スコアの変化[縦軸] (2010年度)

つづいて、やはり表2と同じように入学時の Total スコアを区間に直し、区間ごとにスコア変化の平均とSD、スコアが上がる学生の数・割合を算出し、スコア変化の平均・スコアが上がる学生の割合について、それぞれ平均値と母比率の検定を適用した。その結果を表3に示す。表3からは、(440, 460], (460, 480]の区間でスコアを有意に下げ、その一方でスコアが有意に上がる区間は存在しないことが分かる。

表3：入学時のスコア区分とスコア変化 (2010年度)

入学時のスコア	平均	SD	n	m	m/n
( 0, 400]	30		1	1	1.00
(400, 420]	23.5	33.2	2	1	0.50
(420, 440]	5.7	23.6	26	15	0.58
(440, 460]	- 13.8*	39.1	31	13	0.42
(460, 480]	- 21.1***	34.9	40	10	0.25**
(480, 500]	- 2.13	33.7	15	5	0.33
(500, 550]	- 7.7	46.7	10	5	0.50
(550, 677]	- 3.5	23.3	2	1	0.50

表2と表3から両年度を比較すると、入学時のスコアの平均を含む(2010年度の(460, 480]), もしくはその近傍にいる学生(2009年度の(460, 480], (480, 500], 2010年度の(440, 460])はス

コアを有意に、しかも著しく下げている傾向がみられる。当然のことながらスコアが上がる学生の割合とも強く連動しており、特にこの区間については2～3割程度の学生しかスコアを伸ばせていない。また、平均よりも20点程度低い、440点以下の学生らについては、2010年度については、統計的な有意性は認められないもののスコア変化の平均が上がっており、スコアを伸ばす学生の割合も上述したような区間の学生らよりも高い。このような状況から、大きくスコアを伸ばせていない、むしろ落としているのは、概して入学時の Total スコアの平均周辺の学生、つまり九大生の中では TOEFL スコアという面では平均的な英語力を有した学生らであることが推測される。なお、「標準化テストに見られる九大生の英語能力の変化」（徳見，2010年）によれば、1年生後期時の Total スコアは入学時と大きく変化していないことが明らかとなっており、2年生から3年生以降の期間に、上述したような変化が生じている可能性が高い。

#### 入学時の下位項目スコアから見た変化

前節は Total スコアという全体的なスコア変化について、入学時の Total スコアから見てきたが、次に下位項目についても同様に入学時のスコアからその変化と特徴を確認する。入学時の各下位項目のスコアを、やはり恣意的ではあるものの、35点以下と35点から65点の間は5点刻みでスコアの区間を構成し、各区間に対するスコア変化の平均とSD、スコアが上がる（変化が0より大きい）学生の数とその割合を算出した。平均と割合に対しては、前節までと同じような検定を適用した。2009年度のLC, SWE, VRC に対する結果を、それぞれ表4, 5, 6に示す。なお、n = 0であった区間については表から省いている。

表4：入学時のLCスコア区分とLCスコア変化（2009年度）

入学時のスコア	平均	SD	n	m	m/n
( 0, 35]	0		1	0	0
(35, 40]	3.88***	3.10	17	16	0.94***
(40, 45]	- 0.15	3.67	101	44	0.44
(45, 50]	- 1.10**	3.28	72	22	0.31**
(50, 55]	- 1.50	4.60	8	2	0.25
(60, 65]	- 5.50	4.95	2	0	0

表5：入学時のSWEスコア区分とSWEスコア変化（2009年度）

入学時のスコア	平均	SD	n	m	m/n
( 0, 35]	8.33	9.29	3	3	1
(35, 40]	3.07	6.49	14	10	0.71
(40, 45]	- 0.17	5.15	76	35	0.46
(45, 50]	- 2.07***	4.33	90	21	0.23***
(50, 55]	- 4.35**	6.41	17	2	0.12**
(55, 60]	- 8.00		1	0	0

表 6：入学時の VRC スコア区分と VRC スコア変化 (2009年度)

入学時のスコア	平均	SD	n	m	m/n
( 0, 35]	11.80**	5.67	5	5	1
(35, 40]	5.82***	6.10	17	14	0.82
(40, 45]	2.64***	3.52	55	43	0.78***
(45, 50]	- 0.97*	4.47	87	34	0.39
(50, 55]	- 2.80***	4.49	30	5	0.17***
(55, 60]	- 6.71***	2.93	7	0	0

2009年度については、まず LC と VRC におけるスコア変化の平均が統計的に有意である区間を見ていくと、いずれも各平均 (LC は45.0, VRC は46.4) を境界とした傾向の相違が見られる。平均よりも低い学生はスコアが向上し、高い学生はスコアが下がる傾向が高い。特に VRC については、その傾向が顕著で、意味のある変化の幅であるかどうかは議論の余地がある区間も存在するが、全区間においてスコアの上下動が有意である。その一方で、SWE については、その平均 (45.4) 以上の区間に、スコアが下げる区間が存在するのみである。

次に、2010年度の同様の計算結果について、LC, SWE, VRC をそれぞれ表 7, 8, 9 に示す。

表 7：入学時の LC スコア区分と LC スコア変化 (2010年度)

入学時のスコア	平均	SD	n	m	m/n
( 0, 35]	11.00		1	1	1
(35, 40]	0	4.58	3	1	0.33
(40, 45]	0.64*	2.90	70	38	0.54
(45, 50]	- 1.11*	4.20	45	14	0.31*
(50, 55]	0	5.33	6	3	0.50
(55, 60]	2	0	1	1	1
(60, 65]	1		1	1	1

表 8：入学時の SWE スコア区分と SWE スコア変化 (2010年度)

入学時のスコア	平均	SD	n	m	m/n
( 0, 35]	9		1	1	1
(35, 40]	2.75	3.50	4	3	0.75
(40, 45]	0.09	5.29	58	24	0.41
(45, 50]	- 2.59***	4.71	44	11	0.25**
(50, 55]	- 3.14*	4.50	14	4	0.29
(55, 60]	- 3.60	4.10	5	1	0.20
(60, 65]	1		1	1	1



表 9：入学時の VRC スコア区分と VRC スコア変化 (2010年度)

入学時のスコア	平均	SD	n	m	m/n
( 0, 35]	16.00		1	1	1
(35, 40]	7.50***	2.07	6	6	1*
(40, 45]	1.24	4.73	29	18	0.62
(45, 50]	- 3.82***	5.82	68	13	0.19***
(50, 55]	- 2.30*	5.49	20	6	0.30
(55, 60]	1.00	0	2	2	1
(60, 65]	- 4.00		1	0	0

LC については、2009年度同様、その平均 (45.5) の周辺 ((40, 45]と (45, 50]) でスコア変化に正負の傾向が見られる。SWE も2009年度同様で、基本的には平均 (46.5) を含む区間とそれより高い区間 ((45, 50]と(50, 55]) でスコアを有意に下げることが分かる。VRC については、2009年度とやや異なり、平均を含む区間とそれより高い区間 ((45, 50]と (50, 55]) でスコアを有意に下げ、平均より低い(35, 40]のみでスコアを有意に伸ばすことが分かる。2009年度は全区間で有意な上下動であったが、2010年度はたった 2 区間で、しかも (35, 40]の区間に含まれる学生数はわずか 6 名であり、2010年度 VRC の平均は大きく下がる結果となっている。

下位項目について、入学時のスコアとその変化について見ると、少なくとも次のような傾向が読み取れる。LC と SWE については、両年度で同じ傾向が見られる。すなわち、LC については入学時の LC スコアの平均を含む、またやや高い区間ではスコアを下げ、低い区間ではスコアを上げている。SWE については、入学時の SWE スコアの平均以上の区間で大きくスコアを下げている、平均よりも低いところではさほどスコアを伸ばすことが出来ていない。特に SWE は、表 1 で入学時から有意にスコアを下げていることから、SWE に関しては、入学時には平均もしくは平均よりも高い学生がその項目の実力を維持、さらに伸ばす指導が必要である。また、VRC については、年度によって傾向が大きく変わっている。そこで、両年度の差異に着目してみると、2009年度では存在していた有意にスコアが伸び、なおかつその学生数が多い表 6 の (35, 40]、(40, 45]のような区間が、2010年度で (あえて対応するとすれば表 9 の (40, 45]と考えられるが) 有意にスコアを伸ばすことが出来ていない。LC や SWE のように VRC でも、やはり平均以上の学生らの入学時の実力を維持、伸ばす必要があることに加え、平均未満の学生への積極的な指導の必要性が示唆される。また、「標準化テストに見られる九大生の英語能力の変化」(徳見, 2010年) によれば、1 年生後期の TOEFL 試験では、下位項目スコアは LC と SWE を伸ばし、VRC を下げる傾向が報告されており、3 年生の状況とはかなり異なり、今後のさらなる調査・分析が必要である。

#### 九州大学が求める英語力とカリキュラム

以上の TOEFL スコアの分析を基礎にして、九州大学における英語教育への提案を試みてみたい。まず言語文化研究院が実施している英語教育の現状を示し、その後に指標となる英語能力の基準を模索し、それに対応する英語教育の基本方針を考えてみる。

九州大学の英語教育は言語文化研究院が責任母体であるが、その教育理念は English for Academic Purposes である。English for Academic Purposes とは、学部教育や大学院教育で必要な英文読解能力と英語論文作成へと発展する基礎的英語力を意味する。九州大学の英語教育の目的は、学問的に内容のある事柄を表現し、聞き手（読み手）を納得させる英語能力を身につけさせることである。したがって、流暢な英語を話すことを要求することはない。自分の考える学問的なアイデアを、英語でしっかり伝えることができればよいだけである。だからと言って、英語による発表能力をまったく無視しているわけではない。英語 IIA や英語 IIIA の授業では、学期の最後に英語によるプレゼンテーションを実施しており、伝達能力も成績評価に組み込んでいる。その集大成として、言語文化研究院では、毎年、「外国語プレゼンテーション・コンテスト」を開催して、学生の外国語発表能力の向上を奨励している。

現在の英語カリキュラムは2006年度から導入され、インターネットを利用した英語授業を大胆に取り入れながら（英語 IIB および英語 IIIB）、English for Academic Purposes を推進するために、少人数クラス授業の英語 IIA、英語 IIIA で、英文エッセイ作成と英語によるプレゼンテーションを指導している。その他には、英語 I で英語の文化的背景を理解させるために、言語文化研究院の英語科教員が作成した共通教科書（*A Passage to English*, 九大出版会）を使い授業を実施していたが、2009年度から英語読解能力を九大学生にさらに付けさせるために、読解用のテキストを言語文化研究院英語科が指定して、授業で使用してもらう方法を採用している。このように言語文化研究院英語科では、常に英語カリキュラムの改訂を実施しており、大学生にふさわしい英語を獲得するための英語カリキュラムを提供し続けている。<sup>15)</sup> 2011年度には、現行の英語カリキュラムの改革に着手し、より九大学生に相応しいカリキュラムに変更する委員会を発足させる予定である。しかしながら、どのようにカリキュラムを改訂しようとも、English for Academic Purposes という基本的な方針は堅持されていくであろうと思われる。

それでは具体的にどのような指標を用いて、九州大学で英語教育を実施していけばよいのであろうか。それには CEFR (Common European Framework of Reference) が参考になるとと思われる。Council of Europe のホームページを見ると、CEFR を “Developed through a process of scientific research and wide consultation, this document provides a practical tool for setting clear standards to be attained at successive stages of learning and for evaluating outcomes in an internationally comparable manner” と説明しており、外国語教育の達成度を A1, A2, B1, B2, C1, C2 の 6 つの段階に分類して、それぞれの can-do リストを作成している。CEFR はヨーロッパにおける外国語教育についての指標であるから、それを直に日本の英語教育に移植することは困難と思われるが、ヨーロッパと日本との相違を考慮に入れておけば、日本でも十分に参考にできるものと考えられる。私見では、九大学生はリスト中

<sup>15)</sup> 参考のために、現在実施されている言語文化基礎科目（英語）のカリキュラムを次にあげておく。

	前 期	後 期
1 年次	英語 I 英語 IIA	英語 IIB 英語 IIIA
2 年次	英語 IIIB 英語 IV (文系のみ)	英語 IV

の B2 を目指すべきであると思われる。B2 レベルの全般的な尺度は、以下の通りである。

- A learner can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation.
- A learner can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party.
- A learner can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.<sup>16)</sup>

次に CEFR の「話し言葉」の B2 レベルを記しておこう。

- Range A learner has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much conspicuous searching for words, using some complex sentence forms to do so.
- Accuracy A learner shows a relatively high degree of grammatical control. He/she does not make errors which cause misunderstanding, and can correct most of his/her mistakes.
- Fluency A learner can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he/she searches for patterns and expressions. There are few noticeably long pauses.
- Interaction A learner can initiate discourse, take his/her turn when appropriate and end conversation when he/she needs to, though he/she may not always do this elegantly....
- Coherence A learner can use a limited number of cohesive device to link his/her utterances into clear, coherent discourse, though there may be some 'jumpiness' in a long contribution.

B2 レベルの英語力は、TOEFL ITP に換算すると、567 点から 636 点となる。因みに、2010 年 6 月に実施した 1 年生の標準化テストでは、567 点以上の学生は 17 名しかいない。B2 レベルのような目標を、九州大学で達成するためには、英語授業や授業以外で、英語語彙の習得、リスニング能力の向上、実用的レベルの英文法獲得等の指導が肝要であろう。

語彙の習得に関しては、大学受験時の 3,000 語や 4,000 語では、とても「読み・書き・話し・聞く」ことなどできないので、10,000 語以上程度の語彙を獲得するという地道な努力が必要とされる。<sup>17)</sup>

<sup>16)</sup> 各レベルの基準を知りたい方は、Language Policy Division, Council of Europe のホームページを参照のこと。なお日本では、この CEFR を英語教育に活用しようという動きがある。その代表者である投野由紀夫東京外国語大学教授が、『英語教育』(2010 年 10 月増刊号)に「CEFR 準拠の日本版英語到達指標の策定へ」『英語教育』(大修館、2010 年 10 月)、pp. 60-63 を掲載している。また CEFR の翻訳本、吉島茂他訳編『外国語の学習、教授、評価のためのヨーロッパ共通参照枠』(朝日出版、2004 年)が出版されている。

<sup>17)</sup> 英語語彙については、京都大学英語学術語彙研究グループ編『京大・学術語彙データベース基本英単語 1100』(研究社、2009 年)、東京大学教養学部英語部会編『東大英単』(東京大学出版会、2009 年)、東京工業大学編『東工大英単 科学・技術例文集』(研究社、2011 年)などが役に立つ。特に、最後の『東工大英単 科学・技術例文集』は、理系の学生に必要な英単語と例文が収められている。また『北海道大学英語語彙表』(Hokkaido University English Vocabulary List) がインターネット上に公開されている。なお、本文にある 3000 語、4000 語という表現は、相沢一美他編『JACET8000 英単語』(桐原書店、2005 年)を参考にしたものである。

また日本人英語学習者はリスニングに弱いので、あらゆる機会を利用して（インターネット・映画・テレビ・ラジオ）、リスニング能力向上に努めるべきである。周囲には英語リスニングの機会はいくらでもある。あとは学習者のやる気だけである。前節までの入学時と入学後の TOEFL スコアの分析では、リスニング能力は入学時と3年生以上の成績には変化がないことが証明されている。この事実は、リスニング能力にはまだ伸びる余地が大きいことを示している。

英文法に関して言えば、TOEFL スコアの分析から明らかになったように、SWE のスコアが入学時に比べて低くなっている。これは2年生で全学教育としての英語授業が終了し、3年生、4年生には英語授業がないことが原因として考えられるが、2009年度と2010年度の2年間だけのデータで結論を出すのは早計であろう。これからの分析を慎重に見ていく必要がある。学校英文法は、英語教育者以外の一般人には評判の悪いものであるが、英語を使って世界で活躍している日本人は、英文法の重要性を説いていることを銘記すべきである。<sup>18)</sup> 文法を徹底的に内在化して、それを活用できるまで身に付けなければ、中学校や高等学校で英文法習得に費やした時間は無駄となるであろう。<sup>19)</sup> もちろん瑣末な文法事項に拘泥する必要はないが、曖昧に文法を覚えていては現実の場面では使えない。以上のような指針を念頭に置いて、これからの九州大学における英語教育カリキュラム改革を進めていくべきであろう。

---

<sup>18)</sup> 竹内理『より良い外国語学習を求めて』（松柏社、2003年）、p. 133を参照のこと。

<sup>19)</sup> ただし、日本語以外の言語の文法に触れることは、日本語を客観視できるので、一概に全て無駄とも言えない。外国語学習の利点は、母国語である日本語を客観的に見る機会や視点を提供することである、と筆者らは考えている。英語学習のために、日本語教育が疎かになれば、本末転倒である。