

教育研究方法論にもとづく TOEFL スコアの検証と日本人の英語力について

高 桑 光 徳

「日本人は英語ができない。」テレビや新聞などのメディアを通じて、このフレーズに触れた方は多いであろう。それだけでなく、英語教師や英語学習者からも、同じような発言を聞くことが多い。いわゆる英会話学校や、子どもの英語教室あるいは塾といった英語関連産業の繁栄ぶりをまのあたりにすると、やはり「日本人は英語ができない」と考えている人たちが多そうである。だが、本当に日本人は英語ができないのであろうか。英語ができないと主張する人たちは、その根拠をどこに置いているのであろうか。多くの場合、こうした主張の裏付けには、Test of English as a Foreign Language (TOEFL) と呼ばれるテストのスコアが用いられている。「アジアの諸外国に比べると、日本の TOEFL のスコアおよび順位が非常に低い。だから日本人は英語ができない」という理由付けである。しかし、この主張には、どこか問題がないであろうか。さらに、データの裏付けもなく、勘に頼って「日本人は英語ができない」と結論づける人達もいる。だが、根拠もなく、ただ自分の経験のみで日本人全体の英語力について断言してしまうやり方に問題はないのであろうか。この論文では、TOEFL のスコアの持つ意味を教育研究方法論にもとづきながら検証することを中心に、果たして本当に「日本人は英語ができない」のかを探っていく。

1. 現在の英語教育の置かれた環境

日本人の英語力は本当に低いのか、そしてその裏付けとされる TOEFL のスコアのもつ意味や、勘に頼った理由付けの是非を考えていくにあたって、まず日本における英語教育の現状に触れておこう。

2002年に文部科学省が「『英語が使える日本人』の育成のための戦略構想」を発表したが、この中で英語によるコミュニケーション能力の必要性は以下のように謳われている。

経済・社会等のグローバル化が進展する中、子ども達が21世紀を生き抜くためには、国際的共通語となっている「英語」のコミュニケーション能力を身に付けることが必要であり、このことは、子ども達の将来のためにも、我が国の一層の発展のためにも非常に重要な課題となっている。(文部科学省, 2002a)

そして、この構想を具体化するためのプランとして、中学・高校での達成目標とすべき「国民全体に求められる英語力」が以下のように設定されている。

- ・中学校卒業段階：挨拶や応対等の平易な会

話（同程度の読む・書く・聞く）ができる（卒業者の平均が英検 3 級程度）。

- ・高等学校卒業段階：日常の話題に関する通常の会話（同程度の読む・書く・聞く）ができる（高校卒業者の平均が英検準 2 級～2 級程度）。（文部科学省，2002b）

引用から分かる通り，文部科学省の想定する英語によるコミュニケーション能力とは，英語の会話力と，その会話力と同程度の読む力，書く力，聞く力を指している。言い換えれば，英語によるコミュニケーション能力の育成とは，英語のオーラシー能力（聞く・話す）とリテラシー能力（読む・書く）の両方を育成することが目標となっている。ごく当たり前の内容だが，これが目標となるということは，裏を返せばこれまでの英語教育の方向性に偏りがあったということになる。わざわざ指摘するまでもないが，一般的には日本ではリテラシーに偏った教育がなされてきた。その反省から，オーラシーを含めた総合的なコミュニケーション能力の育成を目指そうとしているのである。これは，言語習得を考えた際に，きわめて常識的な目標であり，伝統的なリテラシー偏重の教育から脱却できそうなチャンスを漸く迎えているといえるであろう。

2. 反「コミュニケーション」主義者と 教育研究における科学的データの 重要性

ところが，この「コミュニケーション」という言葉に明らかに拒絶反応を起こす人達がいるようである。ここではその代表的な例として，鳥飼玖美子・茂木弘道・斎藤兆史の 3 氏を取り上げることにする。この 3 氏を取り上げる理由は，積極的

な出版活動を行いながら，比較的激しい論調で「コミュニケーション」主義の英語教育を批判しつつ，基本的な教育研究の知識が欠けている点において，3 氏に共通点があるからである。しかし，鳥飼・茂木両氏と，斎藤氏には大きな違いもある。鳥飼・茂木両氏は，日本人の英語力が低いことの根拠として TOEFL のスコアを用いているが，斎藤氏は科学的な研究方法がお気に召さないようで，TOEFL のスコアすら使わずに自分の「勘」で論じている点である。斎藤は，英語教育学の成果を「科学的教授法研究の名のもとに行われたアンケートの調査用紙がうずたかく積み上げられただけ」（斎藤，2005, p.21）と評し，また上記文部科学省の改革方針を「英語教育学妄信」（斎藤・斎藤，2004, p.17）と述べるなど，科学的な方法論を用いる英語教育学が相当嫌いなようである。このように，3 者の立場に違いはあるものの，メディアへの露出の大きい彼らの発言は，英語学習者を惑わせる恐れが大きい。以下に順を追って，彼らの主張の問題点を指摘しながら，教育研究における科学的なデータ検証の重要性を見ていくことにする。

3. 日本人の TOEFL のスコアは低いのか

まず鳥飼の主張から見てみよう。

日本人の TOEFL スコアが低いのは，一般に考えられているのとは違い，リスニング力や会話力がないせいではない。通常の TOEFL スコアには，スピーキング力測定は入っていないし，リスニングがよいとは言えないにしても，文法力や読解力だって決してよくない。日本人は文法には強く英語を読むのは問題ないけれど話せない，という前提は，過去のも

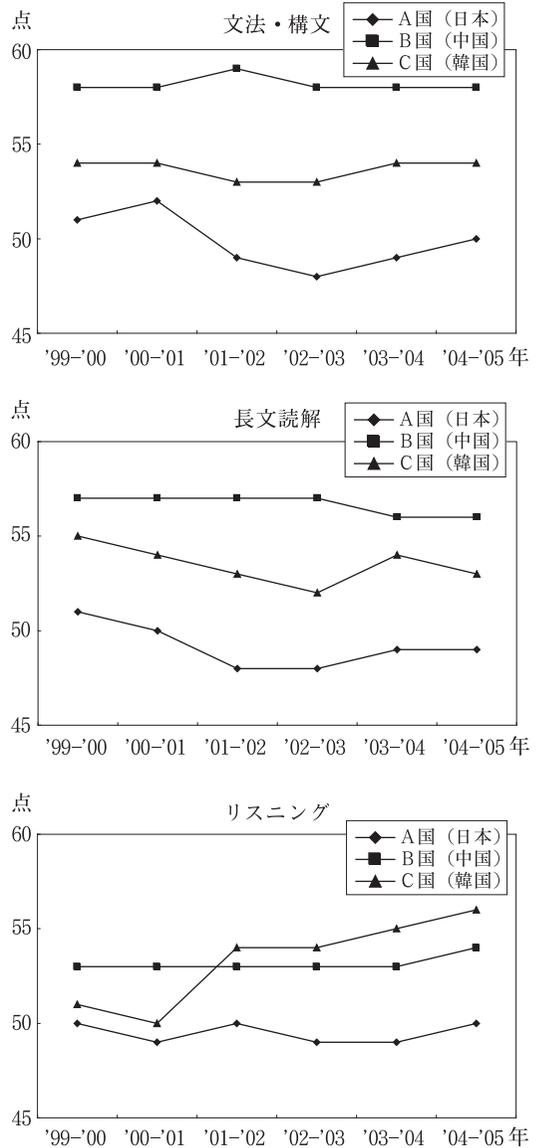
のである。実際は、文法も自慢できるほどではなく、「読む力」が乏しいから、結果として、全体のスコアが低いというのは、スコア分析を見ると一目瞭然である。(鳥飼, 2002, pp.90-91)

このように鳥飼は、まず「日本人の TOEFL スコアが低い」ということを前提とし、その理由を読解力の乏しさであるとしている。鳥飼の主張は、数年後も変わっていないようで、鳥飼(2006)では TOEFL のスコアをグラフ化し、中国・韓国と比べて TOEFL のスコアが低いことを強調している (p.161)。実は、鳥飼のグラフには出典が明記されておらず、データの信憑性も問題となりうるが、TOEFL のデータを公式に発表しているのは、TOEFL を運営しているアメリカの Educational Testing Service (ETS) という組織しかない。したがって、鳥飼もこのデータを用いたと思われる。この ETS が公表しているデータを用いて、鳥飼に出てくるグラフを以下に再現してみる。なお、TOEFL のペーパー版のテストでは、リスニング (Section 1)、文法・構文 (Section 2)、長文読解 (Section 3) (訳語は鳥飼(2006)に準じてある)の順に行われるのでこの順序で取り上げられることが普通であるが、鳥飼が文法・構文、長文読解、リスニングの順に取り上げているので、そのやり方を踏襲することにする。A 国・B 国・C 国という表記についても同様で、鳥飼が本文で国名を最初伏せて紹介しているので、そのやり方に合わせた。

A 国・B 国・C 国を日本・中国・韓国の中から当てさせるというクイズ形式でこの図を示した後で、鳥飼は次のように述べている。

ところが、実際にトーフルの内訳を見ると、

図1 鳥飼(2006)によるセクション別 TOEFL スコアの3国比較



「文法」「読解」がともに、各国に大きく差を開けられているのです。日本は一九七〇年代後半に中国と韓国に追い抜かれて以来、「文法・構文」と「長文読解」で大きく点差を開けられてしまいました。(鳥飼, 2006, p.162)

つまり、鳥飼の主張としては、日本人の英語力は

低く、その原因は文法・構文と長文読解にある、
というものである。ここでは鳥飼を代表に取り上
げているが、後に触れる茂木にも出てくるように、
この論旨の展開は、現在の日本人の英語力が低い
という主張にしばしば使われるものである。

ところが、ここには大きな問題点がいくつか潜
んでいる。もちろん、一番大きな問題点は、ここ
に出てくる TOEFL のスコアから「日本人の英
語力」といった一般化ができないことであるが、
この問題は後で詳述することとし、まずは教育研
究の方法論上で問題となる解釈について見ていき
たい。

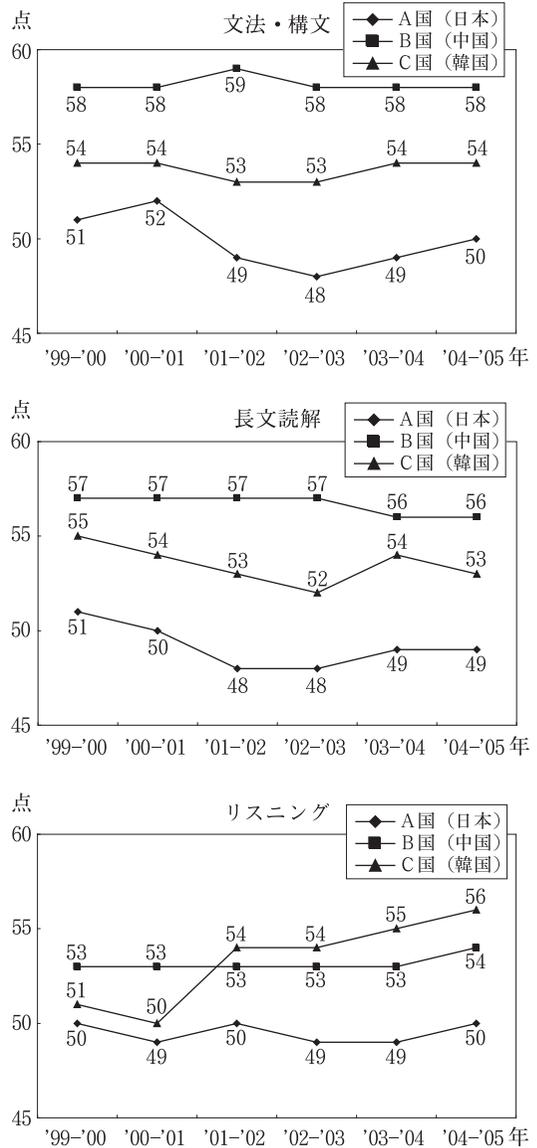
まず、TOEFL のスコアは、他国と比べてその
優劣を決めるものではない。したがって、各国間
のスコアを比較することそのものが不可能である。
このことを理解した上で、鳥飼の論旨の展開の矛
盾を指摘するために、あえて鳥飼の方法に則るこ
とにする。その上で、単純に図 1 のグラフが示す
スコアの差が、鳥飼の言う大差なのかどうかを検
証したい。

確かに、図 1 のような形でグラフを示されれば、
スコアには差があるように感じられる。ではまず、
このグラフに数値を当てはめてみよう (図 2)。

実際のスコアの入ったグラフから分かる通り、
鳥飼の言う大差とは、数点から最大でも 10 点
〔「文法・構文」のグラフにおける 2001 年-2002
年と 2002 年-2003 年の対中国スコア〕であるこ
とが分かる。さて、この差は果たして大きな差で
あろうか。グラフをもう少し変えてみる。ペーパー
版の TOEFL は、コンピュータ版である CBT の
導入に伴い、1998 年 7 月以降は最低点が 31 点で、
最高点がリスニングと文法・構文が 68 点、長文
読解は 67 点となっている。これを踏まえてグラ
フを正しく作り直すと、図 3 のようになる。

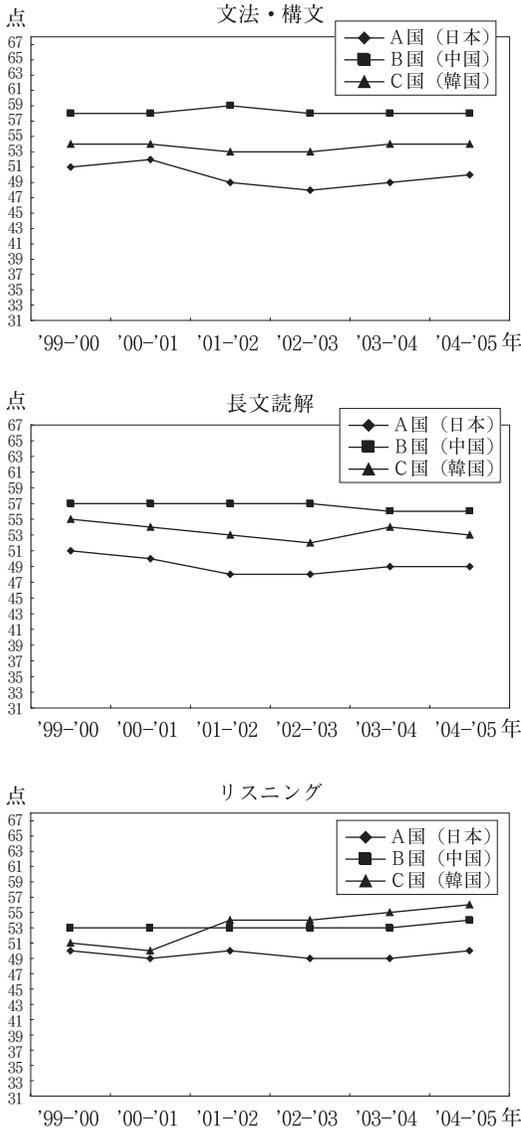
TOEFL の最低点と最高点を考慮すれば、これ

図 2 鳥飼 (2006) を基にしたセクション別 TOEFL
スコアの 3 国比較 (スコア付き)



らが正しいグラフであり、鳥飼が僅かな差を強調
しようとしていたことが分かる。鳥飼自身が、差
を強調したかったことは以下の記述からも分かる。
コンピュータ版の点数ではなくて、敢えてペーパー
版の点数を使用した理由として以下のように述べ
ているのである。

図3 鳥飼 (2006) を修正したセクション別 TOEFL スコアの3国比較



ひとこと、お断り。トーフルは近年、コンピューター・テストに切りかわりつつありますが、この方式は点数の幅がせまく、各国の点差がわかりづらいので、ここでは点数の幅の広いペーパー・テストの点数を使用しています。(鳥飼, 2006, pp. 160-162)

こうした誇張したやり方が、社会科学方法論としては望ましいものではないことは言うまでもない (Gall, Gall, & Borg, 2003; Glass & Hopkins, 1996)。

さて、次に、先程の「数点から10点」という差に話を戻そう。TOEFLにおいて、はたしてこの点差は「大差」なのであろうか。この問題を考えるときにどうしても避けては通れないのは、誤差という考え方である。

日本のように、一発勝負の大学入試が大きな意味をもつような社会にいと忘れがちであるが、通常はテストスコアには多少の測定誤差が含まれるものである。TOEFLに限らず、あるテストで受験者が実際に取ったスコアは、その受験者の真の力を表すスコアに測定誤差を含んだスコアである。したがって、その受験者の真の力がどれくらいであるかを理解するためには、受験者が実際に取ったスコアに測定誤差を考慮した点数域を設定し、受験者の真の力はその点数域のどこかに存在すると考える必要がある。この点数域の設定に用いられるのが標準測定誤差 (standard error of measurement) であり、受験者の真の力を表すスコアとの関係は以下ようになる。

Given the properties of the normal distribution curve... , we can assume that about two-thirds of all test scores will be within $\pm 1s_m$ (plus or minus one standard error of measurement) of the individuals' true scores, and about 95 percent will be within $\pm 2s_m$ (plus or minus two standard errors of measurement) of their true scores. (Gall, Gall, & Borg, 2003, p. 199)

ETS は、数年に一度 TOEFL のマニュアルを

表1 実際の TOEFL のスコアと標準測定誤差を考慮した点数域の例

	文法・構文 (Section 2)	長文読解 (Section 3)	リスニング (Section 1)	合計
実際のスコア	50	50	50	500
標準測定誤差 (s_m)	2.7	2.4	2.0	13.9
$\pm 1s_m$ の点数域	47.3-52.7	47.6-52.4	48.0-52.0	486.1-513.9
$\pm 2s_m$ の点数域	44.6-55.4	45.2-54.8	46.0-54.0	472.2-527.8

出版しているが、この中で測定誤差についても言及している。鳥飼が用いているスコアはペーパー版のものなので、ペーパー版についての最新版マニュアルとなる ETS (1997) の記述を検証する。このマニュアルで ETS は、1995 年 7 月から 1996 年 6 月までの間にアメリカとカナダで受験した受験者のスコアから算出した値として、TOEFL の測定誤差を公表している。これによると、測定誤差は「文法・構文」が 2.7 点、「長文読解」が 2.4 点、「リスニング」が 2.0 点となっている。ちなみに、鳥飼は扱っていないが、合計点についても見てみると、測定誤差は 13.9 点となっている。

もちろん、この測定誤差の値は、上記のように「1995 年 7 月から 1996 年 6 月までの間にアメリカとカナダで受験した受験者のスコアから算出した値」であるため、年度や地域が変われば数値も異なってくる。しかし、ETS が公表しているものでペーパー版について最新のものはこの数値しかないこと、また、TOEFL のスコアは信頼度の数値が高いことから、この数値は他の年度あるいは地域のスコアを検証する際にも、少なくとも目安にはなるものである。

さて、ある受験者がペーパー版の TOEFL で実際に取ったスコアが、「文法・構文」、「長文読解」、「リスニング」のそれぞれのセクションで 50 点だったとしよう。合計点は、3 セクションの平均点の 10 倍で算出されるため、この場合は 500 点となる。ここからその受験者の真の力を表すス

コアが存在する点数域を計算してみると、表 1 のようになる。

表 1 から分かるように、もしある受験者の「文法・構文」のスコアが 50 点であったとすると、その受験者の真の力を表すスコアはおそらく 47.3 点から 52.7 点までの点数域に存在するし、この陳述の確実性をもっと上げるとなると、その点数域を 44.6 点から 55.4 点まで広げる必要があるのである。これが TOEFL スコアの正しい理解の仕方であり、ある受験者の真の力を表すスコアが存在する点数域はそれだけでかなりの幅があることが分かる。

これが、あるスコアと別のスコアを比べるとなると、さらにこの誤差が広がるのである。ETS (1997) によれば、スコアを比較する際には、上記標準測定誤差を約 1.4 倍として考慮する必要があるとしている。したがって、ペーパー版の TOEFL スコアの比較をする際には以下のことを注意すべきである。

One should not conclude that one score represents a significantly higher level of proficiency in English than another score unless there is a difference of at least 39 points between them. In comparing section scores for two persons, the difference should be at least 6 points for Section 1, at least 8 points for Section 2, and at least 7

points for Section 3. (ETS, 1997, p. 31)

鳥飼にしたがってセクションの順序を入れ替えているために分かりにくいですが、まとめると「文法・構文」(Section 2)では最低でも8点、「長文読解」(Section 3)では最低でも7点、「リスニング」(Section 1)では最低でも6点、そして合計点では最低でも39点の差がなければ、あるスコアの表す英語の能力が別のスコアが表すそれよりも高いとは言えないのである。

さて、ここで鳥飼による「文法・構文」「長文読解」がともに、「各国に大きく差を開けられているのです(鳥飼, 2006, p. 162)」という主張に戻ろう。上で述べたように、あるスコアと別のスコアを比較する際には、測定誤差を考慮する必要がある。繰り返し述べているように、各国間のスコアを比較することはできないが、仮に鳥飼の主張するような比較ができたとしても、鳥飼のいうような単純な点数比較はできないのである。もちろん、これまで見てきたのは個々のスコアについてであり、これまでの議論がそのまま受験者全体の平均点に当てはまるわけではない。特にTOEFLの場合、何度でも受験することができるので、1回しか受験しない人もいれば、目標の点数に達するまで受験し続ける人もいる。そもそもそうして算出された「平均点」にどのような意味があるのかということも問題となるが、これについてはまた後ほど述べることにし、ここでは個々の点数と平均点では扱いが異なってくるという指摘に留めたい。これまで述べてきたことをまとめれば、鳥飼がグラフで示そうとした「点数から10点」の差について、少なくとも鳥飼のいう大差がついたという指摘はあたらないのである。

さらに付け加えると、セクション別という鳥飼(2006)のグラフの提示の仕方がかなり恣意的で

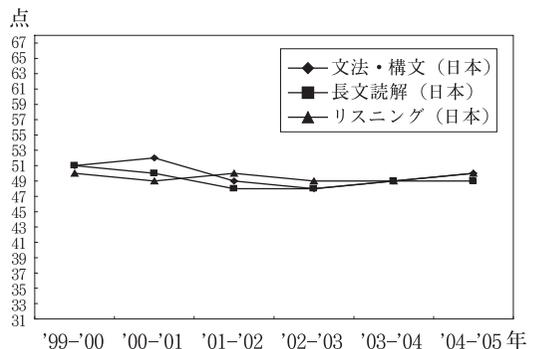
あることが分かる。日本のスコアのうち「文法・構文」「長文読解」が悪いとするならば、他国との比較ではなく、残りの「リスニング」のセクションと比べるべきであろう。3つのセクションを比べた上で、仮にどれかのセクションが著しく低いのであれば、そのセクションの成績が芳しくないと言える。では実際はどうであろうか。同じ鳥飼のデータを、日本だけについてまとめてみる。

図4から明らかなように、日本のスコアについてみると、実は3セクションのスコアはほぼ同じであることが分かる。このことから、鳥飼の主張する「実は『読解力』が不振なのでスコアが低い、という事実」(鳥飼, 2006, p. 158)などどこにもないし、「リーディングや文法のスコアが低いことなどは、ほぼ無視されている」(鳥飼, 2002, p. 101)という主張も事実面に即していないことが分かる。

このように、鳥飼が反コミュニケーションの論陣を張る根拠としているTOEFLの点数は、実は彼女自身の分析とはまったく違う結果を示している。鳥飼自身、以下のようにTOEFL分析の大切さを謳っていることは興味深い。

トータル・スコアを問題にするのなら、その内容を正確に検証してからでないと、英語教

図4 鳥飼(2006)を用いた日本のセクション別TOEFLスコア



育をどの方向へ向かわせるか、という議論は誤った前提で進められることになってしまいます。

数字が意味するものを十分に分析しないまま報道したり、予断に基づいた論評を行ったりすることは、世論を迷わせます。情報を批判的に読み取る能力、いわば「メディア・リテラシー」は、読者だけでなくメディア自身にも求められると痛感します。(鳥飼, 2006, p. 163)

鳥飼も TOEFL のスコア分析の大切さを理解しながら、その方法論が間違ってしまったのである。

なお、ここでは日本人の TOEFL のスコアが低いという見方を鳥飼を代表させて見てきたが、このような考え方は鳥飼だけに限らない。例えば、茂木 (2004) も 1997 年から 1998 年にかけての日本・韓国・中国のスコア比較を、表を用いて行った上で、以下のように述べている。

この表ははなはだショッキングな事実を教えてください。読解力はあるが、会話はできないというのが日本人の英語力であると一般的には思われている。しかしそれはほとんどない大錯覚であることがわかるからである。……要するに、日本の学生が決定的に劣っているのは読解力である、という事実である。(茂木, 2004, pp. 17-18)

また、教育研究者でさえ、こうした誤った見方に陥いる危険性があることは、以下の引用からも明らかである。

日本の大学生や卒業生が標準的英語力において韓国や中国よりかなり劣っているというこ

とは TOEFL (Test of English as a Foreign Language) における成績の差に明らかである。(宮原他, 1997, p. 22)

まさに鳥飼のいうように、TOEFL のスコアの持つ意味を正確に検証できないと、これからの英語教育についての議論は誤った前提で進められる危険性を孕んでいるのである。

4. TOEFL のスコアの順位と英語力

次に、TOEFL の順位から日本人の英語力の低さを論じる方法論の問題点について見てみることにする。日本人の TOEFL のスコアを順位から論じる例は数多く見られる。例えば、小池 (2004) は「日本は TOEFL の点数が 30 年、アジアで最低レベルにある」と述べている。また、読売新聞 (2006) によれば、百ます計算を推奨するなどの陰山メソッドで有名な陰山英男も「留学する学生の英語力を測る『TOEFL』では、日本はアジアで最低だ。日本語と英語は言語的に最も遠い関係にあるからだが、読解が一番悪い」と考えているようである。ちなみに、公表されている TOEFL のスコアから、読解が一番悪いというのは誤った結論であることは上で詳細に見た。このような誤った見方をしている陰山は、教育再生会議のメンバーでもある (首相官邸, 2006)。

このように、TOEFL の順位の低さを日本人の英語力に結びつける論調は多く見られるが、ここでは茂木 (2005) の論旨の展開を代表例として取り上げ、問題点を指摘する。茂木を取り上げる理由は、伝聞形式ではなく、実際に TOEFL のスコアと順位を表で示すことで、少しでも科学的な方法論を用いようとしているが、すでに見た鳥飼の例と同じく、その方法論が誤っているという意

味で、注意を喚起するのに好例だからである。

茂木 (2005) は、1964 年から 2002 年までの日本の TOEFL のスコアとアジアでの順位を掲載している。以下に、茂木の「図表 2 日本の TOEFL スコアの推移」(p. 40) を再現してみる。

大変興味深いことに、鳥飼 (2006) と同様に、茂木 (2005) では資料の出典すら掲載されていない。しかし、まったく同じ表を茂木 (2004) で用いており、そこには「資料出典：国際教育交換協議会日本支部代表部 TOEFL 事業部」(p. 16) と記されている。したがって、この表が、ETS からの委託を受けた TOEFL テストの日本事務局

からの資料であれば、信憑性の高いものであると想像できる。ただし、茂木 (2004) においても、国際教育交換協議会のどの資料からの引用もしくは抜粋であるのかは記されていない。

残念ながら、現在公表されている ETS のデータでは入手不可能な年代が含まれているため茂木が掲載しているすべてのデータを検証するすべはないが、一瞥しただけでもこのデータから英語力を判断することの問題点が浮き上がる。

まず、1995 年から 1996 年より前のデータは、1 年ごとに収集されたものではない。少なくとも 2 年分、場合によっては 5 年分のデータがひとつのデータポイントとして示されている。さらに、2 年ごとのデータが、ある年をまたいで重複している場合もある。データの収集されていない年があることもふまえ、以上のことから、上記のデータから経年比較をすることそのものが問題である。

また、茂木は 1999 年から 2000 年のデータより、ペーパー版の TOEFL のスコアではなく、コンピュータ版の TOEFL のスコアを用いている。ちなみにペーパー版と違い、コンピュータ版の TOEFL は、各セクションの最高点が 30 点で合計点の最高が 300 点となっている。表 2 で 1999 年から 2000 年のデータからスコアが極端に下がっているのはこのせいである。ところで、茂木はなぜ 1999 年から 2000 年のデータよりコンピュータ版のデータを用いたのであろうか。国際教育交換協議会 (2006) によれば、日本でコンピュータ版の TOEFL が実施されたのは、2000 年 10 月からである。もし、茂木の主張するように TOEFL のスコアから日本人の英語力や英語教育について論評が仮にできたとしても、コンピュータ版のデータを用いるならば 2000 年から 2001 年のデータ以降のものにすべきであろう。一方、ETS の資料では、「日本」を母国とする受験者のコンピュー

表 2 茂木 (2005) による「日本の TOEFL スコアの推移」

年-年	総合点	アジア順位 (下から)
1964-66	482	5
1964-67	480	7
1966-71	470	6
1976-77	483	4
1978-80	483	3
1980-82	487	5
1982-84	485	6
1984-86	496	7
1987-89	485	3
1989-91	484	2
1991-93	490	3
1992-94	493	3
1993-95	494	4
1995-96	499	6
1996-97	496	3
1997-98	498	1
1998-99	501	4
1999-00	188	2
2000-01	183	2
2001-02	186	2

タ版のデータが出てくるのは、1998年から1999年の資料からである(ETS, 1999b)。コンピュータ版のデータを初出のものより使いたければ、この1998年から1999年のデータを使うべきであろう。なぜ茂木はそのどちらでもない1999年から2000年のデータからコンピュータ版のスコアを用いたのであろうか。もちろん本人に確かめないと真相は分からないが、先程の資料出典を考えると、国際教育交換協議会のデータをそのまま掲載したのかもしれない。教育研究に携わるものとしては、もとのデータを自ら調査すべきであるが、その手順が踏まれていない可能性がある。

さて、茂木(2005)は、おそらく無意識のうちにとと思われるが、非常に大切な問題提起をしたことになる。日本においてコンピュータ版のTOEFLが導入されたのは2000年10月であるのに、それ以前の「日本」のコンピュータ版のデータが存在することを示したのだ。これはなぜであろうか。繰り返し述べているように、TOEFLのスコアを国別の比較に用いることはできないが、この問題はなぜTOEFLのスコアからある国の英語力について言及することができないかを理解するための鍵のひとつとなる。大切な問題であるので、このことについては、後ほど詳しく考察することにする。

話を順位に戻そう。茂木(2005)は、反コミュニケーションの立場から、彼がいうところの英語力低下について次のように述べている。

コミュニケーション英語教育にその救いを求めて、かじ取りをしてから既に二十数年、JETを導入してから十八年。会話力も含めて学生の英語力は向上していないばかりか、明確に低下傾向を続けています。TOEFLアジア最下位転落はその端的な証標でしょう。(p.37)

そもそも、学生の英語力が明確に低下傾向を続けているという客観的なデータはないのだが、茂木は「TOEFLアジア最下位転落」をその根拠としている。表2で示したかったのは、アジアにおけるTOEFLのスコアの順位が低下傾向にあり、ついに最下位になってしまった、ということであろう。

さて、繰り返し述べているように、TOEFLのスコアで国別の比較をすることはできない。したがって、アジアにおける順位で、日本の英語力を論評すること自体が間違いである。ただ、そう結論付けてしまうと、順位から英語力について考察することの問題点を指摘できない。したがって、ここではあえて茂木の方法論に則った上でその矛盾を指摘していきたい。

順位が下がると、あたかも能力まで落ちたように言われることが多い。スポーツの世界でこの傾向は顕著である。例えば、柔道は日本のお家芸と呼ばれ、オリンピックでメダルを取ることが多かった。ところがメダルの獲得数が落ちてくると、「柔道の力が落ちた」ということになる。柔道のような対人スポーツの場合、勝ち負け以外の客観的な指標がないために、力が落ちたという結論に達することもありえる。ただ、力が落ちたのではなく、他の人の力が伸びたという可能性もある。つまり、順位が下がったことが、イコール力の低下かどうかは、何ともいえない。これが、例えば水泳や陸上のように、タイムという客観的な指標がある場合には、事情が異なってくる。もし、あるオリンピックで平凡なタイムで優勝した人が、4年後のオリンピックでは世界新記録を出したが2位だったとしよう。この場合、順位が下がったから力が低下したということにはならない。明らかに力は上がっているが、それ以上に速い選手が現れたということになるであろう。

TOEFL の場合もまったく同じである。TOEFL のスコアは、もともと順位を比べるためのものではない。例えば、あるクラスに在籍する英語学習者が全員 TOEFL を受験したとしよう。その中のある受験者のスコアが 500 点で、クラスで一番だったとする。数ヶ月後に、またそのクラス全員が TOEFL を受験し、その受験者はスコアを 550 点まで伸ばしたが、順位は一番ではなかったとする。この場合、クラスでの順位は下がったが、測定誤差を考慮しても明らかに最初のスコアよりも高いスコアをとったことになる。したがって、力が下がったことにはならないのである。TOEFL のスコアを解釈する際に大切なのは、その人の相対的な順位ではなく、スコアそのものの変化である。この例の場合、500 点から 550 点にスコアが上昇したことに意味があるのである。

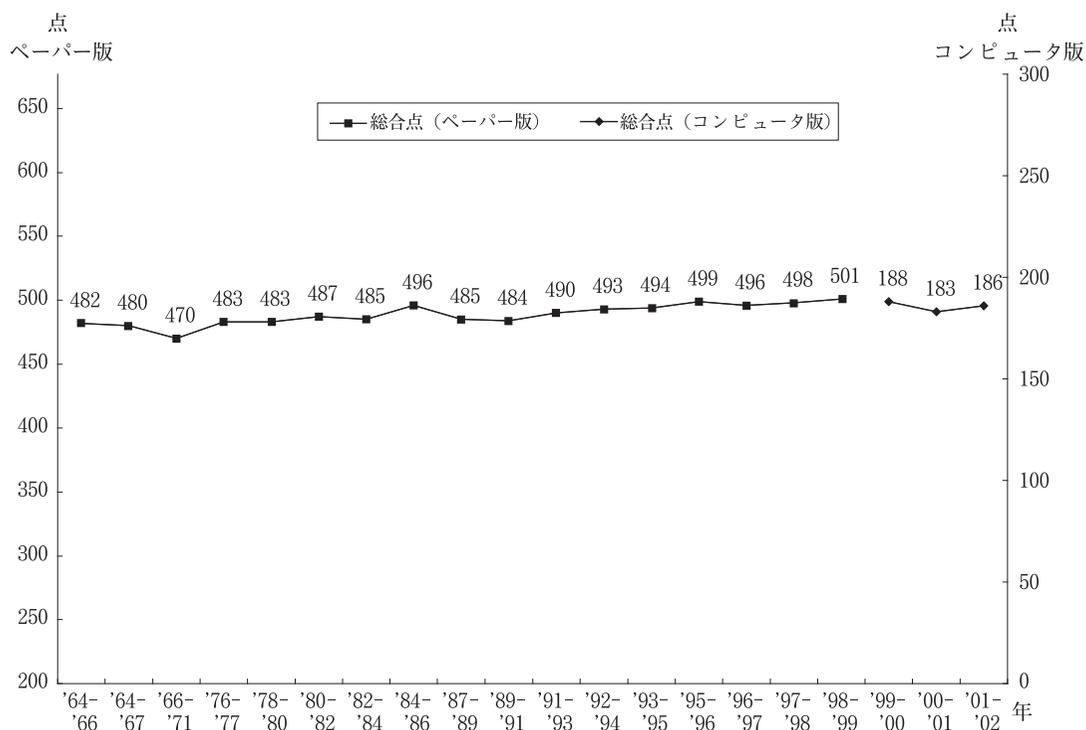
ここでもう一度、茂木（2005）から抜粋した表

2 のデータを見てみよう。ETS が公開しているデータでは TOEFL の経年比較はできないので、このようなグラフを作成することは誤解を与える危険性があるが、あえて茂木の矛盾を指摘するために、表 2 の TOEFL のスコアの推移をグラフ化してみる。

ペーパー版は最低が 200 点（コンピュータ版の導入に伴い 1998 年から 1999 年の年度以降は 310 点）で最高が 677 点、コンピュータ版は最低が 0 点で最高が 300 点となっている。

図 5 を見るとはっきりするが、茂木（2005）に示されている TOEFL のスコアはほぼ横ばいであることが分かる。ペーパー版で 13.9 点（ETS, 1997）、コンピュータ版で 10.8 点（ETS, 2000）という目安になる測定誤差を考慮すれば、実際には差はほとんどないことが分かるであろう。もちろん、公表されている TOEFL スコアの経年比

図 5 茂木（2005）に基づいた日本の TOEFL スコアの推移



較は不可能であるが、仮にそれが可能であったとしても、学生の英語力が「明確に低下傾向を続けて」いて「TOEFL アジア最下位転落はその端的な証標」(p. 37) であるという茂木の主張が誤ったものであることが分かる。そして、この誤った結論は、順位という概念に固執し過ぎたことから生じたのである。

さて、ここでもう一度茂木の主張を見てみよう。順位がアジア最下位になったと主張し、学生の英語力が低下傾向にあると結論づけた上で、次のように述べているのは興味深い。

何も TOEFL で日本が頑張って上に行くべきだなどと言いたいわけではありません。諸外国と比較した数字が明らかに傾向的に下がってきている、という事実を直視せよということです。順位自体はどうでもいいことです。しかしその中身を見るとこれは決してゆるがせにできない重大な問題点が浮かび上がってくるのです。(茂木, 2005, p. 40)

上で詳述したが、茂木の言う通り、「順位自体はどうでもいいこと」なのである。そして、順位についての誤った見方を除くと、TOEFL のスコアが傾向的に下がっているという指摘が思い過ごしであることも分かる。「順位」よりも「その中身」が大切だとしておきながら、その中身を正しく分析する方法論を知らなかったために、鳥飼と同じく茂木も大きく結論を誤ってしまったのである。

5. TOEFL のスコアに基づかない

日本人の英語力批判

ここまで、鳥飼と茂木を例にあげ、TOEFL のスコアのもつ意味を正しく理解しないままに日本

の英語教育批判が行われている危険性を指摘してきた。もちろん、データの誤用は大きな問題であるが、この2人については客観的なデータを用いようとしている点で、まだ改善の余地がある。自らが犯した方法論上の過ちを認めることができれば、正しい方向に議論を進めることができる可能性があるからである。これに対し、英語教育を考える際にかなり危険な存在となりうるのは、科学的データに基づくことなく「自分はこう思う」と言い切ってしまう考え方である。日本で「英語教育の専門家」を自任する研究者や評論家はたくさんいるが、そうした人達の経歴を見ると、「教育」の分野で学位を取得した人が非常に少ないことが分かる。教育は社会科学の一分野であるために、科学的方法論が重要になってくることはすでに見た通りである。ところが、日本の「英語教育の専門家」を多く排出してきた英文学・米文学・言語学といった分野は人文科学の領域であり、伝統的に見て科学的方法論が確立されていない場合が多い。もちろん例外はあるが、人文科学出身の「英語教育の専門家」が科学的なデータに基づくことなく「自分はこう思う」と言い切っている場合が多く見られる。ここではそのようなタイプの典型である斎藤兆史を取り上げたい。

斎藤(2005)は早期英語教育に反対する立場から、日本の英語教育は成果を上げてきていないとし、次のように述べている。

この有名な平泉・渡部論争以降、三十年の間——これはまさしく実用コミュニケーション中心主義が猛威をふるった時代です——日本の英語教育は成果らしい成果を上げていません。ただ科学的教授法研究の名のもとに行われたアンケートの調査用紙がうずたかく積み上げられただけです。四半世紀近く英語教師

をやってきた人間の経験から言えば、日本人の総合的な英語力はむしろ低下しています。「アンケート言語学」の助けを借りずとも、それははっきりと言えることです。(p. 21)

引用から分かる通り、齋藤は科学的教授法研究を重要視していないようである。英語教育学はアンケートだけをやっている訳ではもちろんないし、ましてや「アンケート言語学」などという今のところ確立されていない学問分野と同一視されるべきではない。しかし、齋藤にとって英語教育学はその程度のものであり、自分の経験の方がはるかに優ると考えているのであろう。まさにデータはないが「自分はこう思う」という言い切り型である。

齋藤は、「英語教育の専門家」としてかなりの自信をもっているようで、早期英語教育に反対する理由のひとつとして、次のように述べている。

小学校の英語くらいならだれにでも教えられようという安易な発想があるのかもしれませんが、どんな技芸でも最初の手解が一番難しいのです。そこでいい加減なことを教えたら、あとあと取り返しがつきません。(齋藤, 2005, p. 27)

齋藤は以前にも、同じような早期英語教育反対の流れで「英語教育の専門家でない先生や、日本の言語文化に理解のない母語話者の指導する『英会話ごっこ』だけは、即刻止めなければならない」(齋藤・齋藤, 2004, p. 20) と述べている。ただし、自分は「英語教育の専門家」なので、自分の子供には小学生の頃から英語教育を行っているようである。

小学校六年の私の息子は、小五のときから、毎日一〇分間、『ピーター・ラビット』の暗誦を始めました。最後まで読み終わったら、また最初に戻ってまた読む。それを何度も繰り返すといった学習法で、いまは少しずつ私が文法を教えています。(齋藤・齋藤, 2004, p. 156)

齋藤の専門は英文学でありながら、英語教育学の学位はなくても「四半世紀近く英語教師をやってきた」ので「英語教育の専門家」となるようである。もし、この理屈を受け入れるのであれば、25年間欠かさず読書を続けてきた人は、たとえ文学の専門教育を受けていなくてもみな文学研究者である、ということになるのであろうか。

すでに TOEFL スコアの誤用についての説明で見えてきたように、今のところ、齋藤のいう「日本人の総合的な英語力はむしろ低下しています」(齋藤, 2005, p. 21) という陳述を科学的に裏付けるだけの信憑性をもったデータはない。それにもかかわらず、日本人の英語力が低下していると断言してしまう言い切り方の論調は危険である。齋藤自身、英語教育を改善するためには、「英語が『できる』とはどういうことか、日本人にとって必要な英語力とはどのようなものを十分に議論し、それに基づいた政策を打ち立てる必要があるのである」(齋藤・齋藤, 2004, p. 10) と述べているが、まさにその通りである。だからこそ、そのためにも、客観的なデータに頼らず自分の経験値のみで日本人の総合的な英語力は低下している、という早計な判断に陥ることがないように注意すべきなのである。

6. 「日本人は英語ができない」は本当か

ここまで、反「コミュニケーション」主義を掲げる鳥飼・茂木・斎藤の3氏の論点の矛盾を明らかにしてきた。鳥飼・茂木両氏については、TOEFL のスコアを用いるという、一見科学的な方法論を用いながら、その使い方に誤りがあるために、彼らの主張である日本人の英語力が落ちている、さらにその中でも読解力が落ちているという指摘はあたらないことを見てきた。また、斎藤については、あくまでも自分の勘に頼るという非科学的な論法を展開しているために、日本人の英語力が低下しているという彼の主張は、あくまで推測の域を出ないことが分かった。問題なのは、このような TOEFL の誤用、もしくは非科学的な論拠に基づく「日本人は英語ができない」という主張が、鳥飼・茂木・斎藤の3氏に限られたものではなく、テレビや新聞などのメディアでも広く見られることである。

では、どうしてこのような問題が起きるのであるのか。ひとつには、ある個別のことがらを多くの人々に当てはめる時の一般化の仕方が、本来許される範囲を越えてなされていることが原因となっている。「日本人は英語ができない」という場合、多くの人は「日本人一般」を想定しているようである。日本人全員が英語ができないと考えている人はおそらくほとんどいないであろうから、「日本人は英語ができない」という場合の「日本人」が「日本人全員」を指していることはなさそうである。逆に、「一部の日本人」でもなさそうである。もしそうであれば、「日本人は英語ができない」ことを日本の英語教育の問題と絡めたりしないであろうからである。したがって、「日本人は英語ができない」という場合の「日本人」は、

「いわゆる帰国子女のような海外滞在歴がなく、日本の学校教育で英語を学ぶ、あるいは学んだ平均的な日本人」を指していると思われる。このように定義される日本人が、「日本人は英語ができない」という陳述がなされる時の対象であろう。

ある母集団について何かしらの陳述を行う場合、一番確かな方法は、その母集団全員について調査を行うことである。もし、自分が陳述を行いたい対象となる母集団が小さいければ、全員調査は可能であろう。しかし、多くの場合、「県内在住の大学生」とか「市内在勤の30代の女性」などのように、陳述をしたい対象となる母集団は全員調査が現実的には困難であるほど大きい。この場合、自分が陳述をしたい対象となる母集団から標本を抽出し、その標本に対して調査を行って、その結果から母集団に対しての推論を行うことになる。

抽出を行う際に大事になってくるのが、抽出された標本がもとの母集団の特徴を保持していることである。Glass and Hopkins (1996) の指摘にもあるように、抽出された標本が母集団の典型的な標本でなければ、そこから推論される結論には意味がほとんどなくなってしまう。

A sample is of little inferential value if it is not representative. Many samples reported by the popular media (e.g., volunteer "call in" samples) are essentially useless because they are almost certainly not representative. A sample should be selected in a deliberate fashion from the population so that the characteristics of the population can be estimated with a known margin of error. (Glass & Hopkins, 1996, p. 224)

Glass and Hopkins はさらに、母集団からみて典型的な標本を抽出することがいかに大切であるかを、ただ単に標本の数を増やすことと比べて次のように述べている。

The method used to select the sample is of utmost importance in judging the validity of the inferences made from the sample to the population. The novice is often more concerned with the size of a sample than with its representativeness. A representative sample of 100 may be preferable to a nonrepresentative sample of 1,000,000. (Glass & Hopkins, 1996, p. 225)

ところが、公表されている TOEFL のスコアについては、受験者のべ人数が多いということから、代表性をもつと誤解されるようである。TOEFL についての茂木の引用を見てみよう。

テストの受験者が社会の中で占める位置が国によってある程度異なっている。そのためそれぞれの国の全体水準を同じように反映しているとは言えないからである。しかし、表1に見る通り、日本がアジア最下位に転落した一九九七／九八の日本の受験者数は、一四万六〇〇〇を超えており、韓国一〇万四〇〇〇、中国八万という受験者数の大きさを見ても、かなりの代表性を持つデータであることも否定できない。言ってみれば、かなり本気で英語を勉強している若者の英語力の実態を反映したデータであるということである。(茂木, 2004, p. 14)

前半で、TOEFL のスコアは「それぞれの国の全

体水準を同じように反映しているとは言えない」としておきながら、受験者数の多さから、結局は「かなりの代表性を持つデータであることも否定できない」と結論づけてしまっているのである。

鳥飼や茂木を含め、TOEFL のスコアを用いて日本人の英語力を論じようとする際に陥る落とし穴はここにあるようである。受験者数が多ければ多いほど、その国の「典型的な」スコアだと思いこんでしまっている。ところが、実際はそうではない。まず確認しておきたいが、TOEFL は複数回受験が可能である。上記茂木の引用にもあるように、受験者数が多いことが指摘されるが、これはあくまでのべ人数である。1 回しか受験しない人も入れば、何度も受験する人もいる。もし、TOEFL 受験の目的が、北米の大学もしくは大学院に入学するために必要なスコアを取得することであれば、目標のスコアが取れるまで何度でも受験することになる。逆に、いったん目標点に達したあとは、受験回数は減ってくる。一方、現在の自分の英語力を知りたい人であれば、一度しか受験しないかもしれないし、定期的に受験するかもしれない。要するに、日本人のうちどんな人達が、何回受験しているかは、公開されているデータからはまったく分からないのである。したがって、公開されているデータをもとに「日本人の英語力」について論評することはそもそも不可能なのである。

このことを押しひろげれば、繰り返し述べてきたように、公開されている TOEFL のスコアをもとに、国別のスコアを比較することに何の意味もないことが分かるであろう。母国を日本とする人達のスコアが必ずしも平均的な日本人の力を表すわけではないのと同様に、例えば中国や韓国を母国とする人達のスコアがそれぞれの国の平均的な英語力を表しているわけではない。したがって、

国別のスコアを比較することそのものが誤りなのである。

ETS は、こうした誤解がないように、データの公開をしている冊子の中で、国別の比較をしないように注意を喚起している。ETS はデータを公開する際に、スコアを受験者の「母語別」と「母国別」に分けて報告している。ちなみに、鳥飼にしても茂木にしても、かならずとりあげるのは後者だけである。以下に ETS からの引用を紹介するが、引用の中の「Tables 9 and 10」というのは、それぞれ母語別のスコア一覧と母国別のスコア一覧を指している。

Tables 9 and 10 may be useful in comparing the performance on the TOEFL test of a particular student with that of other students from the same country and with that of students who speak the same language. It is important to point out that the data do not permit the generalization that there are fundamental differences in the ability of the various national and language groups to learn English or in the level of English proficiency they can attain. The tables are based simply on the performance of those examinees native to particular countries and languages who happened to take the TOEFL test. (ETS, 1999 a, p. 5)

上の引用はペーパー版についての記述だが、コンピュータ版が導入されてからも、同様の記述は必ず付してある。つまり、科学的方法論をたとえ知らなかったとしても、データが公開されている冊子にしっかりと目を通せば、公開されているデー

タからは TOEFL のスコアの国別比較はできないことは明らかなのである。

先程の「母語別」と「母国別」の差についても見てみたい。まず、基本的なことがらとして押さえておかなければいけないのは、このデータは受験者の自己申告であるということである。ETS が精査したわけではなく、TOEFL を受験する際の質問事項としてこれらの項目があり、受験者が自分で記入するのである。日本のように、単一民族国家ではないものの、多くの人が日本語を母語とし、日本を母国とする国に長く住んでいると、「母語」と「母国」が密接に関連していると思いきんでしまうのかもしれない。したがって、鳥飼や茂木のように、公開されている TOEFL スコアの「母国別」スコアだけを見て、他国との比較ができると思いきんでしまうのであろう。

例えば、上で引用したように、茂木は「日本がアジア最下位に転落した一九九七／九八の日本の受験者数は、一四万六〇〇〇を超えており、韓国一〇万四〇〇〇、中国八万」（茂木、2004, p. 14）であるとしている。ETS (1999 a) をもとに、この年度の受験者数を詳しく見てみよう。まず「母国別」の受験者数である。

表 3 1997 年から 1998 年の「母国別」TOEFL 受験者数

年 度	日 本	韓 国	中 国
1997-1998	146,439	103,674	79,964

この数値は茂木の主張を裏付けるものとなっている。次に、同じ年度の受験者を「母語別」に表した数値を見よう。

表 4 1997 年から 1998 年の「母語別」TOEFL 受験者数

年 度	日本語	韓国語	中国語
1997-1998	144,579	104,943	161,122

表 3 と表 4 から、ある人の「母語」と「母国」

は必ずしも密接なつながりがあるわけではないことが分かる。日本について見てみると、母国が日本と申告した受験者の方が、母語を日本語と申告した受験者よりものべ 1860 人多い。中国にいたってはその差は歴然としている。ここで、「中国系の人は世界中にちらばっているから中国語を話す人の方が多いのは当たり前であって、母国を中国と申告した受験者から中国という国のスコアを出すことには問題はない」という反論に出会うことは想像に難くない。もちろん 79,964 人の受験者全員が中国国内で英語教育を受けた、あるいは受けているとしても、上で見てきたように、それをもって平均的な中国人のスコアと結論づけることはできない。しかしここで注目したいのは、受験者全員が果たして自分の「母国」と申告した国内で教育を受けているのかどうかということである。

茂木 (2005) による順位を用いた日本の英語教育批判の問題点のところでも少し言及したが、日本でコンピュータ版が導入されたのは 2000 年 10 月からである (国際教育交換協議会, 2006)。ところが、ETS の公表しているデータ上で、最初に日本語を母語および日本を母国とする受験者のスコアが掲載されたのは、1998 年から 1999 年の年度である。それ以降、今のところ最新のデータである 2004 年から 2005 年の年度までのコンピュータ版受験者数の推移をまとめてみると、次のようになる (ETS, 1999b, 2001a, 2001b, 2003a, 2003b, 2004, 2005)。

表 5 を見ると、日本でコンピュータ版が導入された 2000 年 10 月以降のデータが反映される 2000 年から 2001 年の年度から受験者数が大幅に増え、年度始めにはすでにコンピュータ版が受験可能であった 2001 年から 2002 年の年度からは、受験者数が 8 万人規模に達していることが分かる。さて、ここで注目したいのは、日本でコンピュー

表 5 コンピュータ版 TOEFL で母語が日本語および母国が日本であると申告した受験者数

年 度	日本語を「母語」とした受験者数	日本を「母国」とした受験者数
1998-1999	20,495	20,554
1999-2000	21,468	21,636
2000-2001	60,114	60,746
2001-2002	83,357	84,254
2002-2003	81,146	81,749
2003-2004	83,127	83,093
2004-2005	82,235	82,438

タ版が導入される以前の年度についてである。上で見た「母語」と「母国」の差は多少あるものの、1 年でのべ約 2 万人もの人がコンピュータ版を受験したことになる。この時点ではまだ日本でコンピュータ版が導入されていないので、自己申告であるために正確な数字ではなかったとしても、かなりの人が TOEFL を国内ではなく海外で受験したことになる。

これらの受験者についていえるのは、もともと (少なくとも受験の時点で) 海外に居住していたか、あるいはわざわざ受験をしに海外まで行ったことになる。もちろんこれ以上詳しいデータがない以上、どちらが多かったのかは不明であるが、いずれの場合にも、「いわゆる帰国子女のような海外滞在歴がなく、日本の学校教育で英語を学ぶ、あるいは学んだ平均的な日本人」ではなさそうである。前者の場合は、その受験者の英語力が日本国内の英語教育だけで形成されたものとは言えないし、後者の場合は、わざわざ海外まで受験しにいくだけの意欲、そして渡航が自費の場合はさらに経済力があることになり、「平均的な日本人」とはとても言えないのである。

では、日本でコンピュータ版が導入された 2000 年 10 月以降の数字には、上述したような受験者は存在せず、全員が日本国内で受験したので

あろうか。そうかもしれないし、そうでないかもしれない。少なくとも公表されたデータからは何とも言えないのである。大事なことは、ある受験者が自分の母国として申告した国で必ずしも TOEFL を受験しているわけではないし、またその母国で教育を受けたとは限らないということである。

以上のことをまとめると、ETS が注意を喚起しているように、公開されている TOEFL のスコアを国別に比較することはできないし、ましてやそのスコアをもとに、受験者が母国として申告した国の英語教育を論じることは意味をなさないのである。

6. 日本人の英語力とこれからの英語教育

では、一体、「日本人の英語力は低い」のであろうか。正しい答えは、「現状では分からない」である。これまで詳しく見てきたように、公開されている TOEFL のスコアからは「日本人の英語力は低い」という結論は導き出せない。ましてや、鳥飼や茂木に代表されるような「日本人の読解力は低い」という主張はまったく根拠を欠いていることが分かる。一方、斎藤のように、こうした科学的方法論を使わないで同様の結論に達する人達もいるが、これももちろん誤りである。根拠をもち、勘に頼った論が説得力を持たないのはいふまでもない。もし、今後も TOEFL のスコアを使って英語力についての議論をしたいのであれば、まず陳述の対象となる母集団をしっかりと定めるべきである。その上で、その母集団の典型となるような標本を抽出し、データを収集する必要があるだろう。

さらに大切なことは、「日本人」や「英語力」といった言葉の定義付けをしっかりとすることであ

る。英語ができるとかできないといった場合、英語のどういう力のことを想定しているのであろうか。また、日本人すべてが英語の力をつける必要があるのであろうか。あるいは、一部の本当に英語が必要な人達にだけ英語の力をつけさせるべきなのであろうか。今後の英語教育を考える際には、こうしたことがらをしっかりと議論した上で取り組んでいく必要がある。そのためにも、英語教育に携わる人間が科学的な教育研究方法論を身に付けることは必要不可欠である。

現在の英語教育をとりまく状況のところで見たとように、これまでの伝統的なリテラシー偏重の教育方針から舵を切った現在は、まさに英語教育改革の好機といえるであろう。オーラシーとリテラシーのバランスがとれた教育をめざし、総合的なコミュニケーション能力を習得するためには何をすべきかを議論していく必要がある。これからの英語教育を考えるとき、科学的方法論にもとづいて現状をしっかりと認識し、勘に頼るといような予断をもった判断をしないことが大切である。

参考文献

- Educational Testing Service. (1997). *TOEFL test and score manual, 1997 edition*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (1999a). *TOEFL test and score data summary, 1998-99 edition*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (1999b). *TOEFL test and score data summary, 1999-00 edition*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2000). *Computer-based TOEFL score user guide, 2000-2001 edition*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2001a). *TOEFL test and score data summary, 2000-2001 edition*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2001b). *TOEFL test and score data summary, 2001-2002 edition*.

- Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2003a). *TOEFL test and score data summary, 2002-2003 edition*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2003b). *TOEFL test and score data summary, 2002-03 test year data*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2004). *TOEFL test and score data summary, 2003-04 test year data*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2005). *TOEFL test and score data summary, 2004-05 test year data*. Princeton, NJ: Educational Testing Service.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational Research: An introduction* (7th ed.). Boston: Allyn & Bacon.
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Needham Heights, MA: Allyn & Bacon.
- 小池生夫. (2004). 英語 小学校からの一貫教育を. 2004年11月13日. 朝日新聞朝刊. オピニオン2 (14).
- 国際教育交換協議会. (2006). TOEFL メールマガジン. 「日本での TOEFL テスト 40 年を振り返って」. 第9回 (44号). 「コンピュータ版 TOEFL テスト (TOEFL-CBT) 日本上陸」. (<http://www.cieej.or.jp/toefl/mailmagazine/mm44/toefl40th.html>). 2007年1月3日取得.
- 宮原文夫・名本幹雄・山中秀三・村上隆太・木下正義・山本廣基. (1997). このままでよいか大学英語教育 中・韓・日3か国の大学生の英語学力と英語学習実態. 松柏社.
- 茂木弘道. (2004). 文科省が英語を壊す (中公新書ラクレ148). 中央公論新社.
- 茂木弘道. (2005). 小学校英語などとたわごとを言っているときか. 大津由紀雄 (編), 小学校での英語教育は必要ない! 慶應義塾大学出版会. pp. 37-54.
- 文部科学省. (2002a). 「『英語が使える日本人』の育成のための戦略構想」の策定について 英語力・国語力増進プラン. (http://www.mext.go.jp/b_menu/shingi/chousa/shotou/020/sesaku/020702.htm). 2006年11月1日取得.
- 文部科学省. (2002b). 「英語が使える日本人」の育成のための戦略構想 — 英語力・国語力増進プラン —. (http://www.mext.go.jp/b_menu/shingi/chousa/shotou/020/sesaku/020702.htm). 2006年11月1日取得.
- 齋藤孝・斎藤兆史. (2004). 日本語力と英語力 (中公新書ラクレ128). 中央公論新社.
- 斎藤兆史. (2005). 小学校英語必修化の議論にひそむ落とし穴. 大津由紀雄 (編), 小学校での英語教育は必要ない! 慶應義塾大学出版会. pp. 19-36.
- 首相官邸. (2006). 教育再生会議有識者. (<http://www.kantei.go.jp/jp/singi/kyouiku/kousei.html>). 2007年1月2日取得.
- 鳥飼玖美子. (2002). TOEFL・TOEIC と日本人の英語力 (講談社現代新書1605). 講談社.
- 鳥飼玖美子. (2006). 危うし! 小学校英語 (文春新書509). 文藝春秋.
- 読売新聞. (2006). 教師力セミナー in 北海道 優れた授業技術の共有を・特集=北海道. 2006年6月19日. 読売新聞朝刊. 札特1(28).

(2007年1月12日論叢事務局受理)