Bulletin of Faculty of Education, Nagasaki University : Curriculum and Teaching No.46 (2006) 137-151

望ましい英語スピーキングテスト についての予備的考察

―現行英語習熟度テスト・スピーキング部門の分析をもとに―

稲毛　逸郎*・ローン悦子**

(平成17年10月26日受理)

A Preliminary Consideration of English Speaking Tests
― Based on the Analysis of Current English Proficiency Tests ―

Itsuro INAGE *・Etsuko LAWN **

(Received October 26, 2005)

1．はじめに

21世紀を迎え，この地球上で暮らす人間には，限られた地理的・資源的限界の中で互いに「共生」していけるか，また，その目標の実現に対して人間がいかに叡智を絞れるかが問われている。この問に対する答えの重要な鍵の一つが，地球上の人々の相互理解，特に，言葉を用いてのコミュニケーションであることは言うまでもないであろう。

平成10年12月改訂の現行学習指導要領外国語科編（英語）に見られる英語の学力観では，国際化の進展に対応し，国際社会の中に生きるために必要な資質を養うという観点から，特に，英語による実践的コミュニケーション能力の育成や国際理解の基礎を培うことが重視されている。これに伴い，学校教育における英語教育の場においては，この英語を用いての「実践的コミュニケーション能力」をいかに評価するかが盛んに議論されてきている。

英語を用いての「実践的コミュニケーション能力」を評価するということは，基本的には，「聞くこと」「話すこと」「読むこと」「書くこと」の４つの言語技能全般にわたって，様々な現実的場面でのコミュニケーションの達成を目標に，英語をどのように運用できるかという観点から判定する作業を示唆することになるであろう。その中において，特に，英語のスピーキング能力をいかに評価するかは極めて難しい問題となっているのが実情である。

では，一般的に言って，英語のスピーキング能力を測定するということは，何を評価の

*長崎大学教育学部助教授・国際文化講座
**長崎大学大学院教育学研究科教科教育専攻（英語教育専修）

対象にすることになるのであろうか。人間の言葉を用いてのコミュニケーション能力（communicative competence）は，言語能力（linguistic competence），談話能力（discourse competence），社会言語能力（sociolinguistic competence），そして方略的能力（strategic competence）の４つの下位能力からなるという考え方が一般に受け入れられているようである（Canale 1983, Canale & Swain 1980）[1]。しかしながら，異なる文化的背景を持つ人々との，また，異なる言語を話す人々との英語を用いたコミュニケーション活動は，当然のことながら，より広範な要素が複雑に入り組んだ人間の営みであり，外国語を用いてのコミュニケーション能力を構成する要素としては，Canale らが指摘した下位能力に加えて，世界の様々な事象に関する知識・考え（real-world knowledge/thoughts）や，コミュニケーションに臨む態度・姿勢（attitudes）等も含まれると考えるのがより妥当性が高いと言えよう。このように考えると，外国語のスピーキング能力とは，より総合的な能力から成り立つものであり，したがって，外国語のスピーキングテストとは，この総合的な能力を評価できるものでなければならいことになる。

また同時に，スピーキングテストをより客観的なものにするための工夫も必要である。一般的には，望ましいテストとは，(1) 信頼性（reliability）(2) 妥当性（validity）(3) 実施可能性（practicality）の３つの要素をバランスよく具現化しているものであると考えられている[2]。信頼性とは，あるテストが一貫した結果を出す度合い（測定の精度）であり，妥当性とは，あるテストがその狙いとするものをどの程度適切に測定しているかの度合いのことであり，また，実施可能性とは，テストの実施の容易さ，採点・評価作業における容易さの度合いのことである。当然のことながら，望ましい英語のスピーキングテストも，この３要素を備えていることが求められることになるであろう。さらに，松沢（2002）が指摘しているように，英語を用いての「実践的コミュニケーション能力」の評価に関しては，タスク準拠評価（task-based assessment），目標基準準拠評価（criterion-referenced assessment），継続的評価（continuous assessment）という３つの視点に基づいた評価方法のあり方も検討すべきものであろう[3]。

このような背景を踏まえ，本稿では，現時点で制度化され多くの受験者を持つ主要な英語習熟度試験（proficient test）におけるスピーキングテスト（口頭能力試験）について，下記のように大きく３つのカテゴリーに分け，各々の測定方法，評価内容と基準，評価者の角度から比較・検討していくことにする。

1．日本でのみ行われている英語検定試験として，「実用英語技能検定試験」（英検），「国際連合公用語・英語検定試験」（国連英検）の２つを取り上げる[4]。
2．世界的に行われている英語検定試験であり，また日本でも実施されている試験として，TOEIC（LPI），「ケンブリッジ大学英語検定試験」（ケンブリッジ英検），IELTS の３つの試験を取り上げる。
3．日本では行われていない英語検定試験として（イギリスで行われている英語検定試験），TRINITY ESOL EXAMINATIONS を取り上げる。

2．英語スピーキングテストにおける測定方法

まずここでは，各々のスピーキングテストにおいて受験者の発話をどのような形で引き出そうとしているのか，その測定方法（タスクの種類）を次の表にまとめてみる[5]。

表１：６つの英語検定試験におけるタスクの比較

英語検定試験名	タ　ス　ク
「実用英語技能検定試験」（英検）１級～３級	制限インタビュー＋口頭エッセイ（スピーチ）＋制限インタビュー（１級） 制限インタビュー＋情報伝達＋制限インタビュー　（準１級～３級） （情報伝達では，絵を使ったナレーションと英文＋イラストを使用）
「国際連合公用語・英語検定試験」（国連） 特Ａ級，Ａ級，Ｂ級	制限インタビュー＋制限インタビュー＋制限インタビュー （「面接シート」に基づく個人，一般常識，国連に関する質問）
TOEIC（LPI）	自由インタビュー＋自由インタビュー＋ロールプレイ＋自由インタビュー
「ケンブリッジ大学英語検定試験」（ケンブリッジ英検） FCE,CAE,CPE	制限インタビュー＋ロールプレイ＋情報伝達＋制限インタビュー
IELTS	制限インタビュー＋口頭エッセイ＋口頭エッセイ
TRINITY ESOL EXAMINATIONS イニシャル～アドヴァンス	制限インタビュー　（イニシャル） 制限インタビュー＋口頭発表　（エレメンタリー） 制限インタビュー＋インタラクション・タスク＋口頭発表　（インタミーディエット） 制限インタビュー＋インタラクション・タスク＋リスニングタスク＋口頭発表　（アドヴァンス）

馬場（2003：125）は，いくつかの先行研究を参考に，スピーキングテストにおいて複数のタスクを通して口頭能力を測定することの重要性を強調しているが，上記６つの英語検定のタスクを比較してみると，６つのうち５つは２つ以上のタスクを併用する形をとっている。この中で特に Trinity ESOL Examinations は，レベルが上がるにつれタスク数を増やし，上級においては４つのタスクが課されていることがわかる。

また，６つの試験のうち５つが制限インタビューを行っている。そのうち４つの試験は，スピーチ，ナレーション，ロールプレイ等を取り入れることで出来る限り面接委員の一方的な質問に対する応答にならないように，また自然な会話に近い状態になるように配慮がなされている。しかし，日本でのみ行われている英語検定試験の一つである国連英検では，

試験が３ステージにわたって行われているものの，制限インタビューだけで試験が進められる。また，同様に英検も，２級以下では，情報伝達を取り入れてはいるが，面接委員からの一方的な質問になる傾向がある。

６つの試験のうちTOEIC LPIのみ，自由インタビュー形式で試験を行っている。Weir(1990)によると，自由インタビューは，現実生活に近い状況で受験者のパフォーマンスを引き出すことができるが，その反面，インタビューごとにトピックが異なってしまうので，評価基準の一貫性を保つのが難しいという欠点もある。しかし，TOEIC LPIは，その点に関する対策としてフリーカンバセーションの中にロールプレイを折り込み，また評価の客観性を期すために，複数の採点者による評価方式を採用している（馬場（2003：117-119））。

イギリスでのみ試験が行われているTrinity ESOL Examinations は，他のどの英語検定試験にもない特徴が２つある。まず，全てのレベルで，面接中の会話は受験者中心に進められていく点である。次に，受験者が持っている英語力を最大限に面接官に示す機会を与えるために，普段受験者が慣れている（興味がある）トピックについて事前に準備したスピーチを行い，そのトピックに関して討論することが，よりその力を引き出せると判断している点が非常に興味深い。ただしTrinity ESOL Examinations は，TOEIC LPIのように，数名の評価者によって採点されないので，受験者によって異なるトピックから引き出す英語能力の評価基準に一貫性を持たせるのがやや困難であるかもしれない。

この点から考えると，世界的に行われている英語検定試験，あるいは日本以外の国で行われている英語検定試験の方が，日本でのみ行われている英語検定試験と比較すると，よりインタラクティブな（受験者と面接委員との英語を用いての双方向性のコミュニケーションを志向する）傾向にあると言える。日本でのみ行われている英語検定試験のこれからの課題は，一問一答という従来の形式から脱却し，どのレベルでも会話を継続する能力も評価できるようなタスクを準備することにあると言えよう。英検は，その対策の一つとして，2004年第１回の検定試験より二次試験が一部変更されたことに伴い，１級のみ，新評価基準の１つとして「会話を継続する能力」が挙げられている。

３．英語スピーキングの評価の対象について

この節では，各々の英語スピーキングテストにおいて，どのような内容を評価の対象としているかを比較・検討していく[6]。

表２：６つの英語検定試験における試験評価対象の比較表

英語検定試験名	試験評価項目
「実用英語技能検定試験」（英検）１級～３級	①要約力，首尾一貫性　②　会話継続力　③　正確さ，適切さ（文法，語彙）　④　発音（適切なアクセント，イントネーション運用力）⑤　アティチュード（自分の考えや情報を積極的に伝えようとする態度の評価）

「国際連合公用語・英語検定試験」（国連英検） 特A級，A級，B級	① 理解力 ② 会話運営力 （a. 流暢性，b. 発音 c. 構文力，d. 語彙力）③ コミュニケーション能力 ④ 知識
TOEIC（LPI）	① 発音 ② 文法 ③ 語彙 ④ 流暢さ ⑤ 理解力
「ケンブリッジ大学英語検定試験」（ケンブリッジ英検） FCE,CAE,CPE	① 文法，語彙 ② 談話管理 ③ 発音 ④ インターラクティブコミュニケーション
IELTS	① 流暢さ，一貫性，継続性 ② 語彙 ③ 文法，正確さ ④ 発音
TRINITY ESOL EXAMINATIONS エレメンタリー～アドヴァンス	① コミュニカティブスキル ② 文法，語彙，発音 ③ 正確さ（文法，語彙，発音）④ 適切さ（文法，語彙，発音）⑤ 流暢さ（即答）

まず，６つのテスト全てに評価対象（内容）として設定されているのが（各々分類の仕方が多少異なる），文法，語彙，発音の正確さ，適切さであることがこの表からわかる。

次に，「コミュニケーションの能力」として，実用英語技能検定では，会話継続力（2004年より新評価導入）及び首尾一貫性を挙げている。国連英検は，４つの大きな内容の１つとして「コミュニケーション能力」を設定しているが，先にも述べたように，「面接シート」に基づく個人，一般常識，国連に関する質問に対する答え方を評価の対象としている。またその応答を「コミュニケーション能力」と判断している。TOEIC LPI は，「コミュニケーション能力」として会話をスムーズに継続させる流暢さを挙げている。ケンブリッジ英検では，唯一，面接試験が受験者２名で行われるので，ロールプレイ（ペアワーク）などインタラクティブ・コミュニケーションが最も自然な形で測定しやすい試験であると思われる（受験者各々の発言に偏りがないよう，面接官が配分に注意を払いながら試験が進められる）。IELTS は，流暢さ（継続性），首尾一貫性の要素が重要な評価対象になっている。Trinity ESOL Examinations では，コミュニカティブスキルとして，受験者が首尾一貫性のある受験者中心の談話を進め，発話内容を自ら管理すること，また流暢さに関して応答を適切なタイミングですること等が挙げられている。

興味深い点に，特に英検のみが，明確に自分の考えや情報を積極的に伝えようとする態度の評価として，「アティチュード」を評価内容に挙げていることがある。つまり，面接室への入室から退室までの，受験者の「コミュニケーションを図ろうとする意欲・態度」が評価されるのである。同時に，「自分の言葉を理解してもらい，コミュニケーションを持続させようと努めているか」，「音声は明瞭か」，「自然な流れを損なわないスムーズな応答ができているか」などが評価の対象になる。

アティチュードに関する先行研究として，平野（1993）は，「積極的にコミュニケーショ

ンを図ろうとする態度」を「主にコミュニケーション方略を使って積極的にコミュニケーションを図ろうとする態度」であると解釈し，重視すべき評価項目を，「長い沈黙を置かない」，「コミュニケーションを持続しようとする」，「自分から進んで発言（問いかけ，応答）しようとする」「相手と視線を合わせようとする」の４点にまとめている。また，馬場（1994）は，「間違いを気にせずに」「意欲的に」話そうとすることの評価は自己評価以外の方法では不可能であるとする。他方において，母語話者にも声が小さい傾向があったり，感情表現に乏しい人もいることを勘案すると，「声の大きさ」や「感情の込め方」などを態度の評価観点にすることは疑問があることが指摘されている等，「態度評価」については慎重論もある。松畑・石田（1994）は，インタビュー形式のテストを用いて，言語能力を代表するものとして文法能力と，態度が直接関わるものとして方略的能力との関係を調べたが，結果は，文法的能力と方略的能力との間の相関はそれぞれ別のものであると考えている。

最後に，国連英検であるが，国連英検のみが評価の対象に「知識」という観点を採用している。特Ａ級レベルでは，「・・・種々の状況に対応できる英語力ばかりでなく，国際的に通用する知識，情報なども要求される」点に特徴がある。具体的には，「さまざまな国籍の外国人と文化，経済，政治等の多くの分野の問題について，英語で自由に討論できる能力があること，および常識，判断力の点で真の国際人と呼ぶに相応しい水準に達していることが特Ａ級の審査基準です」という記述が見られる。

上記のように評価の対象に関して特徴的なことは，日本でのみ行われている英語検定試験である英検や国連英検は，英語能力以外の内容（アティチュード，知識等）も評価の対象になっているという点である。「アティチュード」が悪いより良い方が望ましいことは明らかであるが，それは基本的には当然のことであって，それを特別に評価の対象に挙げる必要があるのかについては，今後さらに検討されるべきであろう。また，国連英検では，評価の対象として「知識」が挙げられているが，たとえ一般的な会話が流暢にできていても，試験で問われた分野の知識不足のため英語力の評価に影響してしまう点において，その妥当性について議論の余地がありそうである。その点，Trinity ESOL Examinations では，準備したスピーチを討論のトピックにする等，あえて持っていない知識には焦点を当てず，受験者の持っている英語能力を，面接という特別な状態（不自然な状態）でいかに引き出すかということに力を入れている点では興味深い。実際に社会生活をする上で，自分の興味のある分野を英語で流暢に話せるのと同時に，ある程度は，社会問題，時事問題等にも精通していた方がよいのは明らかである。しかし，例えば，ある人が母国語を話している時，ある分野の知識がなかったために，そのことをうまく言葉で表現できなかったとすれば，それ自体が，その人に母国語を話す能力がないとの判断に直接的につながるのであろうか。今後，「英語の総合的な運用能力」を評価することとは何であるのかを再検討し，スピーキング試験の内容を改善していくことが急務である。

４．英語スピーキングテストの評価基準について

この節では，６つの英語検定試験の最高レベルに達成するための各々の基準がどのように設定されているのかについて検討していく。

表３：６つの英語検定試験における最高レベルに達成する基準の比較

英語検定試験名	レベル基準
「実用英語技能検定試験」（英検）１級	大学上級レベル程度　約 10,000 語～ 15,000 語レベル 広く社会生活に必要な英語を十分に理解し，自分の意思を表現できる程度
「国際連合公用語・英語検定試験」（国連）特 A 級	ネィティブスピーカーと同等レベルで，時事問題を含む国内外の問題に関する様々なトピックについて討議ができる程度
TOEIC（LPI） レベル５	スピーキング能力は，教養があり，非常に明晰な英語を話すネィティブスピーカーと同等であり，英語を母国語として使われている国の文化も完全に理解していることが窺がえる
「ケンブリッジ大学英語検定試験」（ケンブリッジ英検)CPE	最上級英語を母国語とする人に劣らない英語力の持ち主であることを証明するテスト。英国や他の英語を公用語とする国の大学で，入学時に必要な英語力の証明として採用されている他，世界各国の機関で認められている
IELTS	全ての状況で，英語をほぼ完璧に使いこなせる程度
TRINITY アドヴァンス グレード 12	ほぼ全ての状況で，英語を容易に理解できる。話し言葉も書き言葉のどの情報も要約でき，議論あるいは説明を非常にまとまりのある形で再構成できる。自然に，流暢に，正確に自己表現ができ，複雑な状況でも微妙な意味の違いを理解できる

それぞれの英語検定試験の最高レベルに達成するための基準として，６つの英語検定のうち５つに「ネィティブスピーカーと同等レベル」あるいは，「英語をほぼ完璧に使いこなせる程度」等の表現が使われている。唯一，英検のみに「広く社会に必要な英語を十分に理解し，自分の意思を表現できる程度」という表現が見られ，この基準は他と比べて明確さに欠け，ややレベルが下である印象を与えている。英検は他のレベルの基準に関しても，その表現がやや曖昧である。例えば準１級では「日常生活や社会生活に必要な英語を理解し，特に口頭で表現できる」，また，２級では「日常生活や職場に必要な英語を理解し，特に口頭で表現できる」等の表現が使われている。

このように，評価基準の表現に関しては，全体的にその基準の表記が曖昧であると言わざるを得ない。例えば，英検１級の基準では，「大学上級レベル程度」「約 10,000 語～ 15,000 語レベル」を挙げているが，近年高校生に求められている学習語彙数が約 2,700 語であることを考慮に入れると，上記の基準は，極端に多い語彙数，つまりネィティブスピーカー程度のレベルを示唆していると考えられる。また，「大学上級レベル」という表現に関してはさらに曖昧であり，少なくとも現時点での一般的な日本人大学生の上級レベルはあまり程度が高いとは言えず，この２つの基準は現実とはかなりの開きがあるという印象を受ける。また，「ネィティブスピーカーと同等レベル」という表記は，レベルとしてはあまりにも広範囲であり，その実態が不明瞭であろう。また，「教養のある」という表記も，

その状態についての基準は，人それぞれ異なるはずであり，基準としては明確であるとは言えない。「英語を母国語としている国の文化を完全に理解していることが窺える」という表記に関しても，「完全」ということはネィティブスピーカーにもあり得ないことであり，評価基準としては不適切であると思われる。

限られた語数で，各々のレベルの基準を明確に表記することは非常に困難であるが，少なくとも誤解されやすい表記は極力避けるべきである。どのような評価基準の表記が受験者にとってわかりやすいか，試験実施機関は今後もより検討を重ね，具体的でかつ現実的な表記をすることが望まれる。

5．英語スピーキングテストの評価者について

この節では，これまで見てきた6つの英語検定試験において，面接委員・レーター（評価官）としてどのような資格が求められているのかを比較・検討していく。

表4：英語検定試験における面接委員，レーター（評価官）比較表

英語検定試験名	面　接　委　員	レーター（評価官）
「実用英語技能検定試験」（英検）	1級 - ネィティブスピーカー1名 準1級～3級 - ネィティブスピーカーあるいは日本人1名(日本人であることの方が多い) 従来は,面接委員用のガイドラインがあり各自責任をもって試験事前に目を通すよう指示があっていたが，近年(2004)，英検は，一部面接委員の研修やトレーニングを始めている(面接委員は，登録制ではあるが，特別な資格は必要ない)	レーターなし
「国際連合公用語・英語検定試験」（国連英検）	特A級，A級，B級 - ネィティブスピーカー1名 面接委員用のガイドラインがあり各自責任をもって試験事前に目を通すよう指示がある。ガイドラインには,発音,文法，流暢さ，会話力（ボディーランゲージを含む)の評価基準が明記されている。また，国連の知識を含めた国際時事の情報も書かれている。面接委員になるために特別な資格を取る必要はない	レーターなし オブザーバー1名

TOEIC（LPI）	インタビューは全てTOEICテストの開発機構である Education Testing Service（ETS）のトレーニングを受け，認定されたネィティブスピーカー１名が行う	レーターは列席しないが，面接は録音され，第２の採点者が再度評価する。評価が異なった場合は，第３の採点者に渡される
「ケンブリッジ大学英語検定試験」（ケンブリッジ英検）	面接試験は２名のネィティブスピーカーによって行われる。実際の対談を進めるのは，１名のみ（対談者兼評価者）。面接委員及び評価者はどちらも，トレーニングを受け，認定されている	１名の評価者が列席するが，一切会話には加わらない
IELTS	ネィティブスピーカーの面接委員１名。面接委員はすべてトレーニングを受け，認定されている。また，面接は面接委員の評価標準化のためすべて録音され，本部で管理されている	レーターなし
TRINITY ESOL EXAMINATIONS	ネィティブスピーカーの面接委員１名。面接委員は，トレーニングを受け，認定されている	レーターなし

まず英検に関しては，１級では，　ネィティブスピーカー１名，準１級～３級では，ネィティブスピーカーあるいは日本人１名であり，これらの級には，面接官が日本人であることの方が多い。従来は，面接委員として英検に登録してある教師に，面接委員用のガイドラインが配布され，それを試験事前に各自責任をもって熟読するよう指示があったようである。しかし2003年，英検はケンブリッジESOLとテストに関する共同研究およびテスト開発を目的とした「ケンブリッジESOLと英検の業務提携についての合意書」に調印し，公正な評価基準およびそれに基づく評価例を作成すると同時に，面接委員の研修やトレーニングを行っている。2005年度の第１回検定より，２次試験形式が一部変更されたが，それに伴い「評価統一のためのトレーニングビデオ」を作成し，英検実施委員，面接委員の研修会を開催している。同年５月９日（日）東京で約600名の教師が参加し，説明会（評価の説明，研修会等）が実施された。

次に国連英検に関しては，２次試験が実施されているどのレベルにおいても，面接はネィティブスピーカーによって行われる。面接委員には，英検同様，ガイドラインが配布され，それを試験事前に各自責任をもって熟読するよう指示がある。ガイドラインには，発音，文法，流暢さ，会話力（ボディーランゲージを含む）の評価基準が明記されている。また，国連の知識を含めた国際時事の情報も書かれている。面接委員は，国際時事，国連の知識を受験者に尋ねる。特に試験で最も重要視されているのは，会話を適切に続けていく能力である。面接委員になるために特別な資格を取る必要はない。

面接場面に列席するオブザーバーは，通例，もと国連大使など国際的に活躍してきた各界の要職にある人物であり，面接試験にオブザーバーとして加わり，英語の能力を評価するのではなく，受験者に国際人としての総合評価を下す役目を担う。

TOEIC（LPI）では，インタビューは全てTOEICテストの開発機構である Education Testing Service（ETS）のトレーニングを受け，そこで認定されたネィティブスピーカーが行う。インタビュアーはフリーカンバセーションを通じて，発音，文法，語彙，流暢さ，理解力などがチェックされる。ただし，これらの項目ごとの評点はあくまでも総合評価のためのプロセスであり，評価は受験者の英語のスピーキング能力に対して総合的に行われる。また，LPIでは評価の客観性を期すために，複数の採点者による評価方式が採用されている。インタビュアーは第１の採点者で，第２の採点者がインタビューの内容をすべて記録した録音テープを聴きながら再度評価する。第１と第２の採点者の評価が異なった場合には，録音テープは第３の採点者に渡され，さらに評価が行われる。

インタビューテストの信頼性を確保するには，インタビュアーの養成やその質の向上・確保（Quality Control）も重要な要素である。そのためTOEIC LPIでは，次のような２つのワークショップ（トレーニングワークショップ）を毎年開催している。まず，新規インタビュアーの養成を目的としたものでは，参加者はすべて英語のネィティブスピーカーであり，開催期間は３日間，専門の講師によるサンプルインタビューや参加者による模擬インタビューなどが行われる。また，最終日にはETS作成のインタビューテープが渡され，参加者はこのテープを採点，さらに，自ら集めた５人分のサンプルインタビューテープとその評価を添えてETSに提出することが求められる。ETSでは提出されたテープを基にインタビューアーの認定を行うが，毎回の合格者は２～３名ときわめて厳しいものである。こうして認定された合格者がTOEIC LPIのインタビュアーとなる。

次に，リフレッシャーワークショップでは，既にインタビュアーとして認定され，実際にTOEIC LPIのインタビューを行っているネィティブスピーカーが対象とされる。このワークショップの目的は，インタビュアーの主観の差，評価のズレをなくし，受験者の能力を常に正しく測定し評価に反映させることである。年１回，実施期間は１日である。ETSではインタビューの方法や評価基準，評価方法の標準化を徹底するため，その詳細をマニュアル化しているが，ワークショップではこのマニュアルに基づき，インタビュアーの質の維持や，評価方法の再確認を行う。TOEIC LPIのインタビュアー資格の有効期限は２年間で，資格を更新するためには，この定期的なワークショップを受けることや課題を提出することが義務づけられている。

ケンブリッジ大学英語検定試験（ケンブリッジ英検）の面接は，ネィティブスピーカーの面接委員（対談者兼評価者）１名とネィティブスピーカーの評価者１名によって行われる。面接試験中は，面接委員のみが２名の受験者（あるいは受験者数が端数の時は３名）と対談する。評価基準は６段階ある。面接委員（対談者）は全体的な評価のみをし，評価者は分析的な５段階（文法，語彙，談話運用，発音，インタラクティブコミュニケーション）を評価する。面接委員も評価者もトレーニングを受け，認定された者である。ケンブリッジ英検は，2002年には，130以上の国々で120万人を超える人々が受験している。評価を標準化するために，ケンブリッジ英検が行われている多くの国々では，面接委員約15名当たりに１名のティームリーダーを決め，他の面接委員のアドヴァイスあるいはサポートをしている。またそのリーダーは，シニアリーダー（ケンブリッジESOLスピーキングテストの代表者）の監督下にある。シニアリーダーは，ケンブリッジESOLから選ばれ，イギリスで行われる年１回の（調整と研修のための）定期会議に出席する。ティームリー

ダーは，シニアリーダーとケンブリッジ英検が行われているそれぞれの国の管理事務局によって選出される。当初のトレーニング後の面接委員及び評価委員の評価標準化は，定期会議と選出されたティームリーダーにより，試験センターへのモニター訪問によって行われる。また，定期会議において，出席した面接委員はスピーキングテストのサンプルを話し合い，ボランティアの受験者を前にして練習面接試験を実施し，評価の一般化と標準化を図る。定期会議で話し合われるサンプルビデオは，経験豊かな評価者グループによって事前に評価されたもので，様々な国民性，異なったレベルの能力を持つ受験者の面接試験が選ばれる。

IELTS の面接委員は，まず適切な教育資格保持者で，テストセンター（120 カ国以上に，約 400 のテストセンターがある）によって採用され，ブリティッシュカウンシルと IDP Education Australia(IELTS Australia) によって認定されている者である。

Trinity ESOL Examinations の面接委員は，全てイギリスを拠点とする ELT のプロである。少なくとも年に１回は，トレーニング，評価標準化，サポートを受ける機会を設けてある。面接委員は，フリーランスの者や高校，大学，言語専門学校に勤務している者，元教師などがいる。また，面接委員は，ネィティブスピーカーである。

面接委員は，日本の英検を除き，どの英語検定試験でもネィティブスピーカーによって行われている。英検のみ，準１級～３級において日本人の教師を面接委員に採用することが多い傾向がある（受験地によりまちまちである)。日本で行われている英語検定試験と，世界的に行われている，あるいは日本以外で行われている英語検定試験の面接委員の大きな違いは，日本でのみ行われている英語検定試験では，従来的には，面接委員は特別なトレーニングを受けておらず（試験の事前にガイドラインを熟読するよう指示を受ける程度)，また認定も受けていない（英語検定指定の資格が必要ない）点である。しかし日本英語検定は，2003 年にケンブリッジ ESOL と業務提携し，2004 年５月には，２次試験改定に伴って（トレーニングの必要性を感じてか）セミナー等を開始している。

スピーキングテスト部門が開始されたばかりの TOEIC LPI も，トレーニングや認定を受けた面接委員によって面接が行われる。またこの TOEIC LPI は，ケンブリッジ英検，IELTS, Trinity ESOL Examinations と同様に，面接委員になるトレーニングも１度きりではなく，面接委員の資格にも期限があるので，定期的に再トレーニングをする必要がある。従って，常にアップデートな評価基準を学習した面接委員によって面接が行われ，面接の標準化（標準化の保ち方は，試験機関により少しずつ異なっている）を保つ努力をしていることになる。

６．今後の検討課題について

本稿では，６つの代表的な英語習熟度テストにおけるスピーキング能力試験を取り上げ，各々の測定方法（タスクの種類)，評価の対象（内容)，評価基準，評価者の角度から比較・検討を行ってきた。その中で，特に英語スピーキングテストの評価者についての比較・検討から浮き彫りになったのは，英語検定試験を実施する機関の多くが，試験官の評価の標準化，また試験官の評価基準を保つことに多くのエネルギーを注いでいるということである。評価者が，特別なトレーニングを受けず，試験実施機関によって作られたガイドライ

ンを試験の事前に読んだだけの状態でスピーキング試験の評価を行うのは，評価の客観性や妥当性が保たれない可能性がある。つまり，スピーキング能力の評価にばらつきが出てくる可能性が非常に高いと考えられる。今後，評価者に関わる問題点，評価者の育成方法，その資格の与え方等に関して，さらに検討を加える必要があるが，これらの点についての考察・検討結果の公表ついては，また稿を改めることにする。

注

１．詳しくは，例えば，Canale(1983)，p.p. 6 -14 を参照されたい。

２．Hughes, A. 2003. Testing for Language Teachers. p.p. 26-50 を参照されたい。

３．「タスク準拠評価」とは，コミュニケーション場面で遂行するタスクを評価課題に設定して，受験者のコミュニケーション能力の達成度を談話レベルで判定する方法であり，「目標基準準拠評価」とは，絶対的評価の視点から，外国語能力の一般的発達過程に照らして，受験者の到達度を示す方法である。また，「継続的評価」は，いわゆる形成的評価の役割を担うものであり，教育期間全般にわたって学習者に関する評価情報を収集し，それらを学習者に対するフィードバック情報として提供することにより，さらなる学習の動機付けや到達感の涵養に役立てようとするものである。

４．簡略化のために，これ以降，「実用英語技能検定」を「英検」，「国際連合公用語・英語検定」を「国連英検」と表記することにする。

５．Weir(1990) は，口頭能力テストを次の８タイプに分類している。

① 口頭エッセイ（verbal essay)：受験者は面接官に対してあるトピックについて話す。

② 口頭発表（oral presentation)：受験者は事前に与えられたトピックについて準備したスピーチを行う。

③ 自由インタビュー（free interview)：あらかじめ話の筋道やトピックなど決められていない面接である。

④ 制限インタビュー（controlled interview）：あらかじめ面接のアウトラインがありそれに従って進める。

⑤ 情報伝達（information transfer)：連続した絵を受験者に見せその描写をさせる。

⑥ 情報伝達（information transfer)：独立した絵について質問応答する。

⑦ インタラクション・タスク（interaction task)：受験者同士がペアを組みタスク遂行のため情報を与え合う。

⑧ ロールプレイ（role play)：受験者同士がペアを組み現実生活を反映させる役割を演じる。

６．Hughes(2003:62) は，CCSE Test of Oral Interaction（Level 3）を例に取り，スピーキングテストが評価対象とすべき内容を次のようにとらえている。

ACCURACY: Pronunciation must be clearly intelligible even if some influence from L1 remain. Grammatical/lexical accuracy is high though grammatical errors which do not impede communication are acceptable.

APPROPRIACY: The use of language must be generally appropriate to function and to context. The intention of the speaker must be clear and unambiguous.

RANGE: A wider range of language must be available to the candidate. Any specific items which cause difficulties can be smoothly substituted or avoided.
FLEXIBILITY: There must be consistent evidence of the ability to 'turn-take' in a conversation and to adapt to new topics or changes of direction.
SIZE: Must be capable of making lengthy and complex contributions where appropriate. Should be able to expand and develop ideas with minimal help from the Interlocutor.

参考文献

国内文献：

小池生夫監修． SLA 研究会編．2003.「第二言語習得研究に基づく最新の英語教育」大修館書店：東京．

杉本薫．2005. 「授業を変える評価」『英語教育』54（4）．pp. 22-25. 大修館書店：東京．

馬場哲生．2003.『英語スピーキング論』河源社：東京．

松沢伸二．2002.『英語教師のための新しい評価法』大修館書店：東京．

松沢伸二．2005. 「こういう評価だけはやめたい！」『英語教育』54(4). pp. 17-19. 大修館書店：東京．

海外文献：

Canale, M. 1983. "From Communicative Competence to Communicative Language Pedagogy." In Richards, J. and Schmidt, R. (eds.) Language and Communication. London: Longman.

Canale, M. and Swain, M. 1980. "Theoretical bases of communicative approaches to second language teaching and testing." Applied Linguistics. 1, 1. p.p. 1-47.

Hughes, A .2003. Testing for Language Teachers. Cambridge University Press.

Van Lier, L. 1989. "Reeling, Writing, drawing, stretching, and fainting in coils: Oral proficiency interview as conversation." TESOL Quarterly 23:3. pp.489-508.

Underhill, N. 1994. Testing Spoken Language. Cambridge University Press.

Weir, C. J. 1990. Communicative Language Testing. London: Prentice Hall.

ウェブサイトからの引用：

英語インタビュー試験専門校テソーラスハウス． 傾向と対策． 国連英検二次試験． 検索 2005/06/20. http://www.mmjp.or.jp/thehouse/home.htm

英語情報ポータルサイト． eigoTown. 英語の資格と検定試験． 検索 2005/06/20. http://www.eigotown.com/eigocollege/exam/exam.shtml

国連英検試験センター（The Unate Test Center). 国連英検の特 A 級について． 検索 2005/08/22. http://www.unate.net/level-sa.html

国連英検試験センター（The Unate Test Center). 国連英検の A 級について． 検索 2005/08/22. http://www.unate.net/level-a.html

国連英検試験センター（The Unate Test Center). 国連英検の B 級について． 検索 2005/08/22.

http://www.unate.net/level-b.html
財団法人 国際ビジネスコミュニーケーション協会. TOEIC(R) LPI について.
検索 2005/05/30. http://www.toeic.or.jp/toeic/lpi/index.html
財団法人 国際ビジネスコミュニーケーション協会. TOEIC(R) LPI について. ワークショップ. 検索 2005/09/16. http://www.toeic.or.jp/toeic/lpi/lpi02.html
財団法人 国際ビジネスコミュニーケーション協会. TOEIC(R) LPI について. 評価基準. 検索 2005/05/30. http://www.toeic.or.jp/toeic/lpi/lpi03.html
財団法人日本英語検定協会. 検定のご案内. 英検とは.「実用英語」とは. 検索 2005/04/30
http://www.eiken.or.jp/info/level/index.html
財団法人日本英語検定協会. 検定のご案内. 各級概要. 検索 2005/06/30.
http://www.eiken.or.jp/info/level/index.html
財団法人日本英語検定協会. 英検審査基準. 検索 2005/06/30.
http://bestesteem.com/eiken.htm
財団法人日本英語検定協会. 検定のご案内. 1級. 検索 2005/07/23.
http://www.eiken.or.jp/info/level/grade-1.html
財団法人日本英語検定協会. 検定のご案内. 準1級. 検索 2005/07/23.
http://www.eiken.or.jp/info/level/grade-p1.html
財団法人日本英語検定協会. 検定のご案内. 2級. 検索 2005/07/23.
http://www.eiken.or.jp/info/level/grade-2.html
財団法人日本英語検定協会 検定のご案内. 準2級. 検索 2005/07/23.
http://www.eiken.or.jp/info/level/grade-p2.html
財団法人日本英語検定協会 検定のご案内. 3級. 検索 2005/07/23.
http://www.eiken.or.jp/info/level/grade-3.html
財団法人日本英語検定協会 ニュース. 英検. 検索 2005/08/29.
http://www.eiken.or.jp/news/backno0802.html
財団法人日本英語検定協会 ニュース. 英検. 検索 2005/09/15.
http://www.eiken.or.jp/news/backno11.html
財団法人日本英語検定協会 ニュース. 英検. 検索 2005/08/30.
http://www.eiken.or.jp/news/backno05.html
ブリティッシュカウンシル. ケンブリッジ英語検定試験5つのレベル. 検索 2005/05/30 .
http://www.britishcouncil.org/jp/japan-exams-cambridge-efl-exams-levels.htm
Ken Howells. "Here are some tips for taking the interview test of UNATE." Retrieved, August 22, 2005, from http://w 1 .nirai.ne.jp/chitchat/kokuren.html
Trinity College London. ESOL Spoken Grade Examinations. Retrieved, August 20. , 2005, from
http://www.trinitycollege.co.uk/index.cfm?fuseaction=esol-spoken.home
Trinity College London. Examiners. Retrieved, September 20, 2005, from
http://www.trinitycollege.co.uk/organizer/induction/gettingstarted/examiners.htm
University of Cambridge. ESOL Examination. Retrieved, June 30, 2005, from
http://www.cambridgeesol.org/support/handbook.html
University of Cambridge. CPE Teaching Resource. Retrieved, August 30, 2005, from

http://www.cambridgeesol.org/teach/cpe/speaking/aboutthepaper/assessment/index.cfm
University of Cambridge. CAE Teaching Resource speaking. Retrieved, August 30, 2005, from http://www.cambridgeesol.org/teach/cae/speaking/aboutthepaper/faqs.cfm