

発表要旨

<第1部>

言語と文化の分かちがたい結びつき—— 異文化対応能力育成のための言語教育

ジョゼフ・ロ・ビアンコ

本論は今回のラウンドテーブルに設定された3テーマにあわせて、スタンダード/カリキュラム、異文化間の言語教育の進展に沿った言語テストおよびアジェンダ設定、オーストラリアにおける最近のカリキュラムやシラバスの改定、について論ずる。

なかでもとくにカリキュラム改定、テスト、全体的なアジェンダ設定について、言語教育に関わる以下のような3つの進展を糸口に論じる。

- ・ビクトリア州主要科目スタンダードの非英語言語のための枠組み(2005)
- ・The Learning Federation of Australia(オーストラリア学習連盟)のアジア言語デジタル学習目標プロジェクト(2003-2005)

そして、全般的なこととして、

- ・オーストラリアにおける最近の文化・言語教育政策の進展、なかでもとくに、アジア言語異文化教育ナショナル・ガイドライン

本論では、こうした進展を例証として、教育制度、教師、コミュニティ全体が義務教育での言語学習の成果として受容でき、評価できるとみなすようになっていくものに、文化についてふつうのことばで表される新思考がどのようなインパクトを与えているのかを考える。とくに、英語が主流言語となっている国において、異文化間の言語学習への関心の高まりのおかげで、学校教育で文化的洞察力や異文化意識が優先され、その結果、測定可能な言語能力への関心が薄らぐという影響がでているかどうか、と問いかける。

アメリカ合衆国のクレア・クラムシュ(Claire Kramsch: 1993)、イギリスのマイケル・バイラム(Michael Byram)、フランスのジュヌヴィエーヴ・ザラト(Genevieve Zarate)などによる画期的な仕事に触発されて、言語教育における文化は急速かつ深い再概念化の過程にある(Lo Bianco and Crozet: 2003)。

ふつうの会話においても、ほかのいかなるレベルにおける言語とその使用においても、言語と文化は互いに関わり合うものである。異文化間の言語教育(Intercultural language teaching)という用語が生まれたのは、コミュニケーション習慣は文化的に固有なものだと思われていることを表し、言語における文化についての情報を学習者に向かってもっと発信せよと教師たちに促すためである。異文化間の言語教授法の台頭は、言語そのものが新たなかたちでイメージされるようになってきていることの反映かもしれない。文化の「表現」に携

わるのは国語の「所有者」たるその国の当局であるのがふつうだった。しかし、グローバリゼーションのもと、言語と国家の結びつきはどんどん緩くなっている。従来は、外国語を教えるための教材、シラバス、各国語の推進をはかる各国の文化機関の活動や言説などが、その対象言語の文化的情報を識別し、推進する「場」となっていた。

結局のところ、ナショナルであるとして、すなわち、ナショナルに限定された独自の文化表象として識別される言語が、もっとも多く教えられてきたのだ。それは、弁別的な、あるいは単一の、民族からなる区切られた存在がネーションだ、とする支配的な概念に基づいていた。しかし、国語とは、まさにコミュニケーション連続体のなかのひとつの部分——いまや言語学や社会学の専門家たちによって「分離した」存在であることをしきりと疑われるようになってきている部分——である。そしてそうした専門家のほとんどが、この分離した存在はたんに政治的に決定されたか、あるいは、政治的に生成された文化的プロセスの結果にすぎない、という結論にいたっている。国家も、グローバリゼーションによって、さらには、文化形成的な慣行やグローバリゼーション——とくにインターネット——がもたらした新たな情報や資料の無制限な流れによって、もっとも厳しい試練にさらされる文化的形成段階にあるのだろう。そうした影響は広範囲に見ることができる。いくつか例を挙げれば、ソマーズの *Bilingual Aesthetics: A New Sentimental Education* (2004) は、グローバルな世界が提供する多言語国家であるがゆえに不可避免的にもたらされる相互作用、混成、過誤、交渉の滑稽さ、トラブル、喜びに光を当てている。強力な国家が台頭してきて以来この方、言語使用が区切られた権威ある国語の束縛から今ほど自由になったことはなかった。ランプトンの「マイクロ相互作用の分析」(2005) は、この混成のひとつの効果を子どものおしゃべりのなかの「クロッシング」、すなわちコード・スイッチングに見る。ランプトンはこれを混合も、切り替えも、過度の混成も望まない権威付けシステムにたいする学習者の一種の「トークバック (talking back)」と解釈する。

こうした進展とともに、多様ではあるが中途半端な言語能力を身につける人が多くなっていることもまた明らかである。わたしたちはみんな、ほんのわずかな時間内に、あるいは特定の状況や領域で、あるいは一定の機能的目的のために、多様な言語に遭遇することが多くなった。その結果が、異なる言語、異なる目的、異なる対話者、異なる場合のあちこちで理解スキルと表現スキルの不均衡が見られるように、たくさんの言語能力が複雑に集まっている状態である。

さらに、外国語学習者とその言語を母語とする人とがコミュニケーションをはかるとき、両者が「第三の場」を創りだす傾向があることもますます認識されるようになってきている。すなわち、両者のダイナミックな相互作用のなかで、コミュニケーション規範が相互作用しているふたつの言語的・文化的な基本システムの混合、妥協、あるいは混成になることが多い、ということである。今ではカリキュラム文書のなかでもこの概念を見かけることがあり、その影響は、どんなスタンダードを策定するのか、スタンダードをどう考え、評価するのか、というこ

とも及んでいるのかもしれない。

本論は、言語の教育と学習におけるスタンダードの理論化についての以上のような批評、傾向、実践などから生じるジレンマを考察する。

参考文献

- Kramsch, C. 1993, *Context and Culture in Language Teaching*. Oxford: Oxford University Press.
- Lo Bianco, J and C. Crozet. 2003, *Teaching Invisible Culture, Classroom Practice and Theory*. Melbourne: Language Australia Publications.
- Rampton, B. 2005, *Crossing: Language & Ethnicity among Adolescents*. 2nd Edition. Manchester: St Jerome Press
- Sommer, D. 2004, *Bilingual Aesthetics, A new sentimental education*. Durham and London: Duke University Press

海外における日本語教育のためのスタンダードとテスト： アメリカからのひとつの見解

リチャード・D・ブレクト

「海外における日本語教育のためのスタンダードとテスト」プロジェクトのような重要な言語イニシアティブに部外者がなにか進言するなど僭越ではあるが、そのような仕事になんらかの経験をもつ世界の政策担当者、教育者、研究者の見解にも見るべき価値は多少あるかもしれない。これまで外国語教育にあまり優れた実績をあげてこなかったアメリカ合衆国だが、グローバルゼーションや2001年9月11日の事件の結果、言語能力の重要性はかつてなかったほど高い認識を得るに至っている。現状では、言語能力のあるアメリカ市民の必要性は広く一般に理解されるようになり、国としての言語能力、とくに連邦政府における言語能力の充実拡大に相当な財源が注ぎこまれている。以下は、こうした観点から申しあげる意見である。

「海外における日本語教育」のための新たなスタンダードとテストが必要なことは明白である。世界各地で日本語教育が人気を得ているためもあるが、それだけではなく、この20年ほどのあいだにさまざまな国で言語のスタンダードやテストが実施されるようになって、スタンダードやテストについての理解が進んだためもある。どのような理解が進んだのか？

1. 第二言語の熟達度を、その言語をどれほど知っているか、ではなく、その言語をどのように使うことができるか、によって定義する。
2. 言語教育がこれまでのように読解中心ではなくなり、コミュニケーション能力が重要な役割を担っている。
3. 言語能力について、「皆無」から「完璧な流暢さ」まで、熟達度を明確に示すスタンダードがある。
4. 言語学習者のある時点での熟達度を有効かつ信頼できるかたちで証明するテスト手続きや方法がある。
5. 集団基準準拠テストから目標基準準拠テストへと移行した。
6. 若年層の学習成果のためのガイドラインがあって、教師や教育管理者に、カリキュラム策定、教材開発と選択、教員育成などについて明確な方向を示す。
7. 学習成果スタンダードに、学習成果の評価を早くから試みるよう規定されている。

以上のような進展は「海外における日本語教育」のためのスタンダードとテストの開発・作成にとって堅固な基盤となる一方で、以下のような重要な問題にとりくんだ前例はほとんどないか、成果が不確実であるために、このプロジェクトがとりくむべきだろう。

8. グローバルな熟達度テストにはどのような機能が意図されているか？ たとえば、共通尺度によって測った言語能力レベルは、職業的な業績評価、学校での進級、などに使われるはずなのか？ それとも、意図しているのは、日本語学習を続けるとか、日本に行くとかいった、学習者の動機づけなのか？ 学習者や学校に向けて、学習、授業、プログラム作成などの向上に有意義なフィードバックをすることで、明確な形成的評価を図るという機能は備えているか？
9. そのスタンダードの基本的な考慮事項は何か？ すなわち、「話す、聞く、読む、書く」なのか、それとも「相互作用、発表、解釈」なのか？
10. そのスタンダードでは段階(5段階?)式スケールを用いるのか、それともスレッシュホールド(しきい)式か？
11. 子ども、青少年、成人の学習成果のあいだにどのような違いをつけるのが適切なのか？ その違いは、全体的な熟達度スタンダードや単一の熟達度テストと、どのように関連するのか？
12. 新しい世界共通テストを行なうことでどのような「波及効果」があり、どのような「波及効果」はないと予想されているのか？ 熟達度とコミュニケーション能力への授業の重点の移行が意図されているのか？ 世界共通テストはどのように誤用されるのか？ 誤用された場合の是正措置はあるのか？
13. そのスタンダードは、どのような文化的価値を伝達するものとして策定されるのか？ 日本にたいすどのような態度を醸成することになるのか？
14. このようなテストを行なうことによって教材、カリキュラム、教師教育などの変化が必要になるが、それをサポートするどのような手だてがあるのか、あるいは今後用意されるのか？
15. 学習者の母語や動機などは世界的にさまざまな違いがあるが、それは世界的な熟達度テストの内容をどのように左右するのか？
16. 策定にどのような配慮をすれば、初等、中等、高等教育のそれぞれで日本語プログラムを効果的に展開していくうえで役にたつのか？
17. 熟達度スタンダードやテストにおいては、方言や言語使用域による違いは考慮の対象となるのか？ それとも、教育のあるネイティブスピーカーの標準的な日本語だけになるのか？
18. 世界的な熟達度テスト実施のさいに、日本語と同系言語を母語とする学習者や、ある程度の日本語のスキルをもって生まれた学習者と、そうした利点のない人たちのあいだに公平性を保つためには、どうすればいいのか？
19. 強い承認のもとに世界中で実施されるテストの安全保障については、どのような問題が生じてくるのか？

20. このようなスタンダードとテストのセットを、国際交流基金による投資はもちろん、主催国の特権も認識されるかたちで、世界中に広めるにはどうすればいいのか？

プロジェクト推進の過程で、このほかにも多くの疑問が提起されるにちがいないが、ここに挙げた疑問でさえ、満足いくかたちでの対応は不可能である。それでも本論では、こうした疑問や問題点を、ラウンドテーブルでの討論に、ひいては、「世界における日本語スタンダード」の開発・策定にも、役だつように提示する所存である。

韓国の言語政策におけるスタンダードについて： 学習指導要領と大学修学能力試験外国語テスト基準を中心に

李 徳奉(イー トクボン)

1) はじめに

韓国の言語政策から読み取ることのできるスタンダードを視野に入れた言語観は、中等教育の学習指導要領により定まるものであり、大学修学能力試験外国語テスト基準などにその具体像がよく表れている。本稿では、学習指導要領における学習目標と、大学入試における出題基準を紹介することにより、韓国における言語観を概観したい。

2) 学習指導要領

韓国の言語政策的特長は、中等教育の学習指導要領によく表れている。中等教育では、英語をはじめ日本語、中国語、ドイツ語、フランス語、スペイン語、ロシア語、アラビア語などの外国語が教えられている。現在適用されている第7次学習指導要領に表れている言語教育の目標としては、全ての言語に共通している点として、コミュニケーション能力を大事にしていることが挙げられる。とりわけ、コミュニケーション機能中心の外国語教育が強調され、流暢さが大事にされている。

意思疎通機能は、言語4技能を中心に述べられているが、中でも、言語行動の理解が目立つ。意思疎通機能の材料についても、挨拶機能、情報伝達機能、要求の機能、態度の伝達機能、談話の展開機能などを中心に具体的な機能が挙げられている。

また、英語以外の言語教育においては、異文化理解教育がより強調されており、中でも、日本語の場合は、異文化に接する際の態度の教育に力を入れている。いわゆる、交流のための外国語教育に徹しているといえる。学習指導要領における言語観の具体像は、大学修学能力試験に詳しく表れている。

3) 修学能力試験とは

韓国の4年制大学に進学するためには、全国レベルの共通試験である「修学能力試験(略称:修能試験)」の成績を提出しなければならない。それぞれの大学では、もっぱらこの成績によるか、論述試験を加える形で合格を決めている。したがって、この「修能試験」の成績は大学入試の決め手になるわけである。この試験に、2001年から日本語等のいわゆる第二外国語科目の試験が復活し加わることになり、そのための出題指針の研究報告書が1999年12月にまとまった。以降の試験は、この指針に従って出題されている。

4) 意思疎通能力とは

意思疎通能力とは、言語体系に対する知識と、これを実際の意思疎通の状況において活用する能力を指す。言語体系には、音韻論、形態論、統辞論などの言語規則であり、規則の活用は語用論が相当する。語用論の範疇には、社会言語学的、社会文化的能力、談話的能力、戦略的能力等が相当する。

5) 意思疎通能力の出題指針

「修能試験」の意思疎通能力評価の要素は、Lussier (1992,p.53)に基づいている。ルシエによると、意思疎通能力には、言語学的能力、社会言語学的・社会文化的能力、談話的能力、戦略的能力等が含まれる。このような評価目標を達成するためには、実際のコミュニケーション場面において問題を解決する課題中心のテストが望ましい。しかし、大学入試という制約的な状況を考えると、間接的な評価方法を取るしかない。したがって、外国語能力のテストには、外国語発音の識別能力、語彙力、文法理解力、コミュニケーション機能の理解や活用能力、異文化理解能力などが含まれる。

6) 評価目標二元分類表

評価目標二元分類表は、内容領域と行動領域からなるが、その項目の構成は次のとおりである。

(1) 内容領域

- ①発音と表記法(10%)、②語彙力(10%)、③文法理解力(10%)、
- ④コミュニケーション機能の理解と活用能力(60%)、⑤文化理解能力(10%)

(2) 行動領域

- ①聞く、②読む、③話す、④書く

それぞれの能力は、知識、理解、応用に分かれる。出題率は、聞く・話す能力40%、読む能力50%、書く能力10%からなる。

7) レベルの設定

言語レベルについては、現在のところ、「1」と「2」に分けて設けられているが、「1」は、入門や初級に相当するものであり、「2」は、中級に相当する。ただし、学習語彙数以外は、その基準はなく曖昧なままである。というのは、中等教育における学習レベルが初級レベルに止まっているからだと思われる。

8) 今後の課題

今後の課題は、言語教育のスタンダードを確立することや、それぞれの言語観別特徴を生かすことを中心に研究が行なわれている。

参考文献

Lussier, D.(1992) Evaluer les apprentissages, Hachette

韓国教育課程評価院研究陣(1999) 大学修学能力試験第二外国語領域出題指針開発研究、韓国教育課程評価院

欧州評議会の言語能力スタンダード

ジョアナ・パンティエ

欧州評議会の尺度に使われている能力記述文は、Swiss National Science Research Council（スイス国立科学研究機関）が1993年から1996年にかけて実施したプロジェクトの成果を基盤として策定された。このプロジェクトは、1991年に欧州評議会が開いた、言語の学習、教授、評価における一貫性と透明性をめぐるシンポジウムのフォローアップ作業として行なわれたものである。その目的は、*Common European Framework of Reference for Languages (CEFR)*の能力記述スキームのさまざまなパラメータによる能力尺度を策定することにあった。

この能力記述スキームによれば、言語の使用および学習のいかなる形態も以下のように記述できる、とされている。

言語の使用というとき、言語学習をも包括して考える。これは人によって遂行される行為の一部である。人は個人としてまた社会的存在として一連の能力(competence)を持っているが、それには一般的能力(general competence)と、特別なものとして、コミュニケーション言語能力(communicative language competences)の二者がある。そして、各自が利用できる能力を使いながら、さまざまなコンテキスト(context)で、さまざまな条件(conditions)下で、さまざまな制約(constraints)の下に言語活動(language activities)に携わる。その際テキスト(texts)を産出するか、あるいは受容するという言語処理(language process)に携わることになる。そこで作られるテキストは特定の生活領域(domains)に属するテーマ(themes)と関連する。またその際課題(tasks)の成就を目指して最も有効と思える方略(strategies)を使う。こうした行為を当事者自らが観察・モニターする中で、上述の能力はそれぞれ強化されたり、修正されたりするのである(翻訳注:この一節は吉島・大橋他訳:2005による)。

言語の使用と学習の複雑さを考えれば明らかだが、いかなるコミュニケーション言語活動にもそれぞれ独自の性格があり、一般的な言語能力を規定するのはきわめて困難な仕事である。

ALTE (Association of Language Testers in Europe)のメンバーが作成した用語集によれば、言語能力は、ある言語についての知識とそれを使用するためのスキルの程度と一致する。しかし、CEFRの能力記述スキームにおいては、ある言語について高い知識とその使用上の高度なスキルを有していても、状況、条件、制約によってある特定の瞬間における能力が左右されたり、テーマや生活領域によって変化したりする可能性のあることが明確にされている。学習や使用のための方略の開発も、人によって、また、異なるパラメータによって、

違ってくるだろう。

しかし、言語の学習、教授、評価における一貫性と透明性を向上させるためには共通スタンダードが重要である。そのためにCEFRは言語能力のための共通の尺度と能力記述を行なうためのスキームを提供することを目的としている。

尺度はそれぞれの言語能力によってもたらされる成果に応じて設定された6段階の能力レベルから成る。レベル数を6と決めたのは、異なる領域での上達を明示しながら、無理のない範囲で首尾一貫した区別もするために適切な数だからである。しかし、地域のニーズに対応しながら共通システムとも合致させるためには、より細かなレベルやカテゴリーを規定する必要があり、「ハイパーテキスト」から枝分かれするアプローチを提案している。

CEFRは、3つの「メタカテゴリー」——コミュニケーション活動、方略、コミュニケーション言語能力——のための尺度表も記述文一覧に掲載している。

これらの「can-do statements」は、Swiss National Science Research Councilのプロジェクトにおいて、国際的な公共領域で使われている言語熟達の尺度や1993年に欧州評議会の関係筋から入手できた言語熟達の尺度を綿密に分析した結果、直感的、量的、質的の3つの手法を組み合わせ、一定の基準で尺度化された。「スターティング・ポイント」となった特定のレベルはなかった。

その後これらの尺度はALTEによって、学習者1万人による自己評価アンケート、および、一連のアンカー項目を使った測定が行なわれ、その妥当性が確認された。ALTE と CEFR の記述文のあいだの相関度は0.97だった。

さらに、CEFR に基づいた自己診断を目的とする言語評価システム、DIALANGによる妥当性の確認も行なわれた。測定の結果、信頼性係数が0.899であるという結果もでた。

CEFRの言語熟達尺度はあくまで例示的なものであり、さらに手を加えて進化させるべき余地を残している。また、(CEFRに基づいた) European Language Portfolioのための記述文バンクが用意されていて、これは一般に公開されて誰でも自由に使うことができる。

さまざまなコミュニケーション言語スキル(聞く、読む、口頭での産出、口頭でのやり取り、書く)のための尺度が存在すれば、モジュール方式コースの構成のためにも、市民の複言語能力の開発のためにも役にたち、明確な学習目標を設定することも達成度を評価することもできる。

現在欧州評議会は*Relating Language Examinations to the CEFR* (CEFR関連言語試験)のマニュアルのパイロット版を刊行しており、それに付随する例示資料集も開発中である。

参考文献

Common European Framework of Reference for Languages: Learning, teaching, assessment,
Cambridge University Press, ISBN 0 521 80313 6

吉島茂/大橋理枝他訳『外国語の学習、教授、評価のためのヨーロッパ共通参照枠』朝日出版社(上記図書の日本語訳)

〈第2部〉

リフレクティブ・オンライン評価と経験的教育

カール・フォルスグラフ

教師も、学習者も、教育内容と成果についての明確なスタンダードがあってはじめて具体的な教育・学習目標を設定することができる。さらにスタンダードがあることで、学習成果を厳密に測定し、試験結果について意義のある解釈を下すこともできる。本論では、オンライン試験が生み出すフィードバックの循環をとりあげる。すなわち、オンライン試験によって成績データが教師や学習者にリフレクト(反映)され、その結果として、評価が教授内容をつくり、それによって学習者もよりよい評価達成に向けて備えができる、という積極的な循環がいかにして可能になるのかを考察する。

従来の試験の特徴は、情報の流れが学習者から試験者へと一方的なことである。たとえば、現在行なわれている日本語能力試験では、受験者は、質問にたいする解答というかたちで膨大な量の情報を発信する。しかし返されるのは、1か、2か、3か、4という、たったひとつの数字だけだ。ところがオンライン試験では、成績データがつねに中核サーバーに保存されているために、学習者、教師、研究者、学習者の親など、誰でもアクセスし、分析することができるのが、その大きな利点のひとつである。関係者に成績評価情報を提供するこのプロセスがリフレクティブ・オンライン評価と呼ばれている。

しかし学習者、教師、管理者などに評価やデータについての基本的理解がなければ、このデータも正確には解釈できない。したがって、教師や管理者に高度な評価リテラシーがあることと、試験設計者が情報をはっきりと提示して結果の意義ある解釈ができるように教師や管理者を教育することが不可欠である。

教育者に評価リテラシーがあれば経験的教育ができる。すなわち、学習者の成績データに応じてカリキュラムや指導法を調整するプロセスが可能になるのである。教師が経験的教育を実践すれば、カリキュラム内容を学習者にインプットするのではなく学習者の成果に焦点を置いた、学習者中心の授業ができる。この意味で、スタンダードと評価は学習者中心の教育のカギとなる。

本論では、スタンダードに基づいたリフレクティブ・オンライン評価の一例として、Standards-based Measurement of Proficiency (STAMP)(スタンダードを中心とした熟達度測定)をとりあげる。とくに、学習者と教師に成績データを報告する通知メカニズムを重点的にとりあげる。さらに、アメリカの教師たちがこうしたデータ報告をどのように利用して、指導法の向上、学習者の親とのコミュニケーション、言語プログラムの推奨を行なっているかも考察する。

日本語能力試験にとっても潜在的な意義は大きい。日本語能力試験も、明確なスタンダードが策定されれば、受験者を大まかにレベル分けするという現行の機能を越えて、微妙なニュアンスでの診断を可能にするデータをプログラムそのものにも、個々の教師にも提供できるだろう。評価だけではなく、日本語教育そのものにも著しい効果があるかもしれない。

成功を目指してランゲージズ・ラダーをのぼる 外国語のための国家認定制度

ケイト・グリーン

2002年12月に発表されたイギリスの「国家言語戦略」は、言語能力と異文化理解はもはや「随意のオプション」ではなく、市民であるために不可欠な要素であることを明確にしている。言語能力は、経済活動にとって重要なものというまでもないが、職業的将来性の幅を拡げたり、国内でも、広く世界においても、人びとの相互理解を深め、文化やことばの障壁を壊してグローバルな市民意識を育てたりと、わたしたちの生活のじつに多くの側面で決定的に重要な働きをする。

「国家言語戦略」の中心にあるのは、既存の資格認定の枠組みを補完する選択制の認定システムを導入して、国民の言語スキルを証明しようという意図である。外国語の能力を証明するために作られるこの新しい認定スキームは「ランゲージズ・ラダー(言語学習のはしご)」と呼ばれる。学習者はひとつ以上の外国語のひとつ以上のスキル(聞く、話す、読む、書く)について「はしご」を一段ずつ上がっていくのである。

6つのステージ(National Qualifications Framework のEntry レベルから Level 8 までと、Common European Framework のA1から C2 までに対応する)については外部評価が行なわれる。各ステージ内の各スキルについてはcan-do statementsが用意され、担当教師はそれを形成的評価として裏書する形で使うことができる。

2006年までには、20以上の言語について最初の3ステージ、すなわちBreakthrough、Preliminary、Intermediateの外部評価の利用ができるようになる。Advanced、Proficiency、Mastery の3ステージの導入と、さらなる言語の提供に向けた開発プログラムは現在進行中である。

「ランゲージズ・ラダー」の主な特徴を以下に挙げる。

- ・各スキルの各グレードにcan-do statementsがある
- ・4つのスキルがそれぞれ個別に評価される
- ・学習者はひとつ以上の言語のひとつ以上のスキルについて評価を受け、ステージを上がっていくことができる
- ・ステージとステージの境界については外部評価がある
- ・オンライン及びいつでも必要な時に受験できる方式の試験が開発される
- ・各ステージ内の各グレードについて担当教師が評価することもある
- ・国および国際的な既存の資格レベルと対応している

can-do statementsは、ひとつ以上の言語のひとつ以上のスキルについて、その上達と熟

達度をモニターするために、学習者も教師も使うことができる。

教師は、教師評価モデルを利用して、教室用教材や各スキルの各グレード用の can-do statements に関連した中核的なタスクセットを使い、学習者の上達をモニターし、グレード・アワードを認めることができる。受容的なスキルの課題はすべて事前テスト済みである。この認定スキームを採用したい教師は、研修に参加して、各レベルで要求されているスタンダードを熟知し、認識できるようになり、Accredited Teacher (有資格教師)として認定されなければならない。

各スキルの外部試験の問題は、各ステージの全グレード用の can-do statements に対応している。受容的なスキルの問題はすべて事前にテストされ、産出的なスキルの問題は試行中である。現在はフランス語、ドイツ語、スペイン語で Breakthrough、Preliminary、Intermediate の各ステージの外部資格試験が、初等、中等、成人教育部門において、試験的に実施されている。2005年5月には全国規模での実施が開始され、2005/6年度には中国語、フランス語、ドイツ語、イタリア語、日本語、パンジャブ語、スペイン語、ウルドゥー語の8つの初期導入言語の評価ができるようになるだろう。

試験実施中のこの資格認定については、現在、言語機能および広範なコンテキスト領域に関連した内容規定があり、国が定める中等教育のための現代外国語カリキュラムに対応している。この内容規定は実験段階終了時に再検討されるだろう。現在は初等教育部門の試験と、中等教育と成人教育の共通試験とが区別されており、さらに成人学習者専用の教材の開発も進められている。

余暇と観光旅行、健康と社会的ケア、ビジネスなど、職業に直接結びついたコンテキストの試験が開発される可能性もある。

AP (アドバンスト・プレースメント) の外国語テスト

デイヴィッド・バウム

AP (アドバンスト・プレースメント) プログラムとは、主としてアメリカ合衆国で行なわれている、中等教育機関とカレッジや大学との共同の教育的取り組みである。1955年の発足以来、意欲ある高校生に高校に居ながらにして大学レベルのコースを履修する機会を提供してきた。また、外国語など多くの科目に共通する信頼性の高い評価法をはじめ、カリキュラムについての的確なガイドラインも提示してきた。外国語については、現在フランス語、ドイツ語、ラテン語、スペイン語のコースが実施されているが、近い将来は、中国語、日本語、イタリア語、ロシア語のコースと試験が新設されることになっている。これらの新科目は Evidence-Centered Design (根拠を中心としたデザイン) モデルにしたがって企画プロセスが進められている。これによってそれぞれのAPコースの目標が——ひいてはそれぞれの試験から引きだされる根拠が——対応するカレッジ・コースのカリキュラムと調整され、中等教育レベルでも適切に実施できるものになる。

新しいコースと試験の細目は、カレッジや大学の教職員と高等学校レベルの熟練教師で構成されるタスクフォースによって決定される。中国語および中国文化、日本語および日本文化、イタリア語およびイタリア文化のタスクフォースは、それぞれ独自に、コース内容をカレッジのカリキュラムと結びつけ、さらにはその試験と Standards for Foreign Language Learning (外国語学習スタンダード) との調整もはかりながら、作業を進めている。

Standards for Foreign Language Learning (外国語学習スタンダード) は、1990年代後半に合衆国内の語学教師の組織の連合体によって開発されたもので、その後、カリキュラム策定権限をもつ各州の教育省によってさまざまなかたちで実施されるようになった。この「スタンダード」は、コミュニケーション、文化、関係、比較、コミュニティの5つを重点的な目標分野としているが、新たなAP外国語テストには、スタンダードの構成基本原理となるこれらの目標分野のなかでもコミュニケーションと文化についての目標が明確に反映されている。

ドイツ語試験の最良実施(ベストプラクティス)—— ゲーテ・インスティトゥートの場合

カタリーナ・フォン・ルックテッセル

ゲーテ・インスティトゥートはドイツ共和国の文化機関である。世界的に活動を展開し、世界の80以上に及ぶ国に付属機関を置いている。その主要な目的のひとつが、海外でドイツ語教育を奨励することをとおして、ドイツと各国とのあいだに文化的交流を促進することである。ドイツの海外文化および教育に関する政策を担当しているゲーテ・インスティトゥートが、語学コースの運営や語学試験の実施をその使命のひとつとしていることは驚きかもしれない。それは、政府の仕事というより、明らかに商業ビジネスと思われるからだ。じつは、ゲーテ・インスティトゥートがドイツ語試験を開発し、市場供給するには3つの理由がある。

- 1) ゲーテ・インスティトゥートが語学コースを提供する目的は、ドイツ語を教えることだけではなく、外国語教育のための最先端モデルを提供することにある。ドイツ語コースを実施すれば、その成果をはかるための試験が必要になる。これまで世界の多くの国々で、ドイツ語コースを実施する最初で唯一の機関がゲーテ・インスティトゥートだった。そのため、試験の開発でも当然ながら最初であり、唯一であった。そして、ゲーテ・インスティトゥートの試験はすぐさま国際的に広く知られ、認知される場所となり、最終的にはドイツ語の知識を公式に証明するものとなった。興味深い事実としては、最初に公式に実施された試験が基礎レベルではなく、現在ならCEFRのC2レベル、あるいはC2+レベルに相当する試験だったことを挙げておきたい。
- 2) ゲーテ・インスティトゥートは語学教育機関としてばかりではなく、研究開発の面でも先進的であり、教育・訓練のための教材のほかにも、必要とされるさまざまなレベルの試験を作成している。現在はCEFRを有し、ALTEの一員でもあり、すべての認定をそれらのA1からC2までのレベルに対応するかたちで実施している。
- 3) ここ10年ほど政府予算が一貫して削減されており、ゲーテ・インスティトゥートとしても、作成した製品の販売にますます力を入れて、コストをカバーする必要に迫られてきた。そうしなければ、活動規模を大幅に縮小しなければならないからだ。言語能力認定が大きな市場になっていることには驚かされる。どうやらグローバル化が進む現代世界においては、特別な知識をもっているだけでは十分ではなく、特別な資格をもっていることを認定によって証明しなければならないようだ。人びとが言語を学ぶだけではなく、検定試験に合格してその能力を証明する書類を手に入れたがるのも無理はないのかもしれない。

質的なスタンダードを維持するかぎり、ゲーテ・インスティトゥートの商業的意欲がその公的
使命と衝突するとは思わない。したがって、言語能力認定について語ることは、すなわち、
その開発、実施、販売のあらゆる分野におけるクオリティ・コントロール・システムについて語
ることになる。それにはゲーテ・インスティトゥートが過去半世紀にわたって獲得してきたよう
な評価や経験が必要とされるだろう。今回のラウンドテーブルの討論では、ゲーテ・インス
ティトゥートがどのような質的なスタンダードを維持しているか、この分野でどのように活動して
いるのか、という内容でお話することになる。

フランス文部省による外国語としてのフランス語認定： DELF・DALF、およびTCF

ブリュノ・メーグル

フランス文部省は、フランス語を母語としない成人および十代の人たちを対象に、外国語としてのフランス語についてのあらゆる種類の公式認定を提供しており、世界160カ国以上で承認されている。Centre International d'Études Pédagogiques（国立国際教育研究所、略称CIEP）は文部省の国際行政機関で、そうした認定の開発、運営、商業取引の責任を負っている。

2002年以降CIEPは、認定機関としてAssociation of Language Testers in Europe（ALTE）会員としても認められている。

外国語としてのフランス語（FLE）能力にたいしてフランス国内と国外で提供している認定には2種類ある。

- ・ ディプロマ
 - DELF (Diplôme d'Études en Langue Française) フランス語認定ディプロマ
 - DALF (Diplôme Approfondi en Langue Française) 上級フランス語認定ディプロマ
- ・ 試験
 - TCF (Test de Connaissance du Français) フランス語能力試験

DELF・DALFの2つのディプロマとTCF試験はフランス文部省による公式の認定であり、フランス国内の全大学に承認されているのはもちろん、ベルギーとスイスの教育担当省庁にも認められている。

現在フランス文部省の提供している外国語としてのフランス語能力認定（DELF・DALF、TCF）は、すべてCommon European Framework of Reference for Languages（CEFR）およびALTEのスタンダードと完全に合致している。

1. DELF と DALF

1985年以降、DELF・DALFは次の3つのレベルで認証されている。

- ・ DELF第1レベル—4つの試験に合格した受験者に発行されるディプロマ。
- ・ DELF第2レベル—さらに2つの追加試験に合格しなければならない。
- ・ DALF—さらに4つの追加試験に合格しなければならない。

上位レベルの受験に期限はない。

1985年以降に10の評価科目をひとつ以上受験した人の数は2,700,000人（最近の4年間の

年平均は350,000人)にのぼる。

しかし2005年9月1日からはシステムが変更され、ディプロマのレベルが6つに分けられる。以下のように、DELFが4つのディプロマに、DALFが2つのディプロマに分かれ、その結果、CEFRとALTEとに完全に対応するかたちになる。

- ・ DELF A1
- ・ DELF A2
- ・ DELF B1
- ・ DELF B2
- ・ DALF C1
- ・ DALF C2

それぞれの試験の内容もEuropean Framework(ヨーロッパ共通言語参照枠)のcan-do statementsに正確に対応して、以下のようなになる。

- ・ 各試験で4つのスキルが評価の対象になる。
- ・ 受験者は能力に応じてどのレベルでも受験できる。
- ・ 試験で50%以上の点数を取った受験者が合格し、ディプロマを取得する。
- ・ 認定に有効期限はない。

DELFの初級レベルであるA1とA2、B1とB2では、成人用と十代用の2種類の異なる試験が行なわれる。

DELFもDALFも、世界960カ所(フランス国内60カ所、海外900カ所)にある試験会場で受験できる。海外でのDELFとDALF試験はフランス大使館(文化部)によって実施される。試験会場として通常は以下の施設が使われる——Instituts Français、Centres Culturels Français、Alliances Françaises、および、現地大学のフランス語科。

2. TCF®

TCFはCIEPが2002年に発足させた。以来70,000人が受験している。また、発足当初年に国際的な品質認定であるISO 9001, version 2000を取得した。

TCF受験者は試験結果によって、言語学習、教授、評価のためのCommon European Framework of Reference for Language(ヨーロッパ共通言語参照枠)に規定されている6レベルのどれかに分類される。TCFは信頼できる、厳正な公式試験で、すべての項目は事前にテスト・分析されて初めて妥当と認められ、基準化される(試験に使われるのは事前テスト済の項目だけである)。

TCFの必須セクションでは実施が容易な多項式選択問題が80問(聴解力30問、言語構造20問、読解力30問)出題される。

必須試験(1時間30分) — 聴解、言語構造、読解

選択試験—話す(15分)、書く(1時間45分)

採点はCIEPによって集中管理され、およそ2週間で成績が出て、成績票が発行される。

- ・ 成績票の表面には、点数と全体的な言語レベル、さらに、受験した言語スキルごとの can-do statementsが記載される。
- ・ 成績票の裏面には、欧州評議会の6段階の言語能力レベルを総合的かつ詳細に説明した記載がある。

TCFの有効期限は2年である。これは、フランス語を母語とせず、なんらかの理由で DALFディプロマをもたない学生のためにフランス文部省が実施する唯一の公式試験である。そのため、フランスの大学に出願する外国人学生はすべて受験しなければならない。

DELFとDALFと同じく、TCFについても、試験は各国のフランス大使館、Instituts Français、Centres Culturels Français、Alliances Françaises、および現地大学のフランス語科を会場に行なわれる(およそ150カ国にはほぼ500カ所の会場がある)。フランス国内にある大学はすべて正式な試験会場になる。

漢語水平考試(HSK)のスタンダード

謝 小慶(シェ シャオチン)

中国の漢語水平考試(略称HSK)は、外国人、海外在住中国人、中国内の少数民族など、中国語を母語としない人びとの中国語能力認定を目的として企画・開発され、標準化された国家試験である。

HSKは基礎、初等、中等、高等の4つのレベルの漢語水平考試で構成されている。高等レベルでは口答試験と筆記試験の両方が行なわれる。漢語水平考試は中国国内と諸外国において毎年定期的実施され、一定以上の成績をおさめた受験者には漢語水平証書が発行される。HSK試験は現在のところ、国内33都市の59会場と、海外33カ国の92の会場で実施されている。2004年には92,995人の外国人が受験したが、そのうち32,423人は国外の会場で、60,572人は国内で受験した。さらに、国内の少数民族も120,000人が漢語水平考試を受験した。

HSKは、アメリカ合衆国のACTFLガイドラインとヨーロッパのALTE/CEFRをモデルとし、さらに、さまざまな調査に基づき、独自のスタンダードを確立したか、あるいは、確立に向けた作業を進めているところである。以下にその調査研究項目を挙げる。

1. 『中国語能力のための漢字と単語のレベル別シラバス』を制作・刊行した。そこでは、漢字2,905字と単語8,822語を、もっとも易しいAからもっとも難しいDまでの4レベルに分類した(漢字については、Aが809字、Bが804字、Cが601字、Dが700字、単語はAが1,033語、Bが2,018語、Cが2,202語、Dが3,569語となっている)。レベル分け作業は、大規模コーパスによる客観的頻度統計と、経験豊かな学者や中国語教師による主観的な判断とを組み合わせて行なわれた。このシラバスはHSKの発展にきわめて重要な役割を果たしており、随時見直しが行なわれている。第一版が1992年に刊行され、2001年には改訂版が出版された。2003年には中国国内の少数民族向けの改訂版が刊行された。
2. 『中国語能力のためのレベル別文法シラバス』を制作・刊行した。1,168の文法項目を難易度別にAからDの4レベルに分類しており、Aは129項目、Bは123項目、Cは400項目、Dは516項目である。
3. 『中国語能力のためのレベル別コミュニケーション機能シラバス』については制作中である。すでにおおまかな言語機能リストは固まっていて、リーディングおよびリスニングによる情報収集、評価、ナレーション、説得、感情表現、思考の方略的表現など42項目が含まれている。
4. HSK試験の有効性検証とともに、成績をより詳細に解釈する方法も探っている。

TOEICの*Can-do Guide*を参考に、経験的調査と研究を実施した。たとえば、HSK受験者のコミュニケーション能力について*Can-do Guide*の自己評価アンケートを使って調査した。

5. HSK試験の重要な機能のひとつに、中国の大学への入学志願者の中国語能力判定がある。理工系に入学するにはHSK3級の能力が最低基準であり、文科系には6級の能力が必要である。HSKでは、Angoff法、境界線グループ法、グループ対照法などさまざまな経験的基準設定法を用いて、妥当な合否ラインを設定している。
6. 中国語学習者の目的に合わせたスタンダード確立に向けた作業を進めている。HSK(ビジネス用)、HSK(秘書用)、HSK(観光用)など、新たな試験群を開発する一方、こうした新たな試験のための新たな語彙やコミュニケーション機能リストを作成している。
7. HSKの高等レベル試験には、口答と筆記の2部門があり、それぞれの試験について、明確に規定された基準にしたがって合否の判定をしなければならない。口答試験についても、筆記試験についても、ACTFLおよびALTEの基準を参考にして、評点のための詳細にわたる基準を設定している。能力レベルの異なる受験者の典型的な解答例を注意深く選びだして、評価基準として用いている。
8. 試験基準の確立もたしかに難しいが、その基準を安定的に維持することも同じくらい難しい。基準を維持する上で重要な手法が等化である。われわれは大きな努力をはらってテスト得点の等化をはかってきた。
9. 10年以上にわたる努力の積み重ねの結果、1998年は、項目応答理論(IRT)にもとづき、コンピューター試験問題作成システムに支えられたHSK 項目バンクが稼動を始めた。項目バンクからコンピューターで自動的に試験フォームを作成する場合に重要になるのが、試験の言語的側面のコントロールである。HSKでは、統計的パラメータに加えて、試験内容のバランスにも配慮し、項目バンクに蓄積されている項目ごとに、評価の対象となる言語スキル、知識の内容、関連する言語機能、の3つの言語パラメータを規定した。

日本語能力試験の課題

松井 嘉和

日本語能力試験は、広い受験者層を対象にした大規模一斉試験であり、実施に際しては、問題の妥当性や信頼性の確保という観点からの自己点検および情報公開という観点からの透明性の確保を方針としてきた。

年1回全世界で同時に行なう日本語能力試験は、1984年に約7,000人の受験者で開始して以来、順調に受験者数が増加し、21回目の2004年には、海外39の国・地域、99都市において243,242人、国内18都市で58,956人、合計302,198人を数えている。

この試験は、4つのレベル(1級～4級)、3つの類別(文字・語彙/読解・文法/聴解)からなっており、受験者は、各類と合計の素点を通知される。また、合格者(1級70%以上、2～4級60%以上の得点を得た者)には認定書が交付される。

この試験は、情報公開を行ない、透明性を保つことに努めてきた。シラバスは、「出題基準」の名によって、問題作成者の手引きとして1994年に出版された。それによって、教育の一定の基準とされ、現場に様々な影響を及ぼすこととなった。また、試験問題は、初期の一時期(1985年から1989年まで)を除き、出版物の形で公開されている。さらに、検証作業として、初年度から『分析評価に関する報告書』がまとめられていて、項目分析など様々な観点からの分析結果が記載され、妥当性、信頼性が検討されている。この報告書は、1990年からは一般に公開され、日本語試験の研究のための資料にもなっている。

既に、第一回の試験を分析した1984年の『分析評価に関する報告書』以来、様々な課題が指摘されて来たが、このほど、最新の学術的成果を踏まえて、現代社会の様々な場面で必要となる課題遂行能力とそのためのコミュニケーション能力を測定するための試験へと改定する作業が、2004年に始まった。2008年には改定後の新試験を実施することを計画している。

20年ほど前は、初・中・上級のレベルの暗黙の共通理解つまり経験上のイメージとしてのスタンダードが日本語教師の間に存在し、それに基づいて1級～4級の基準が設けられたが、それが通用しなくなってきた。それは、従来のスタンダードが専ら大学の予備教育に基づいていたため、学習者の多様化に対応できなくなったためである。

そこで、多様化した受験者が自分の日本語能力で具体的に何ができるかを確認し、以後の学習に資することができるよう、課題遂行能力の観点から、can-do statementsの形で記述する方向で検討を行なっている。すなわち、積み上げ(項目列挙)から「課題遂行能力」へと基準の転換を探ることとなった。先行事例としてCEFRをはじめ他の言語の能力基準の事

例を調査中であるが、日本語能力を測る上では漢字の存在が大きな検討課題となると思われる。試験科目も再編することを検討中であり、口頭能力試験、作文試験についても実施の可能性を検討する。得点表示についても、尺度得点表示を検討している。現在は予備テストを使用した事前等化(pre-equating)は行なっていないが、モニター試験などにより、新しい試験では何らかの形で等化を図ることを検討している。