

言語テスト開発過程の記述と検証

—実践知の共有をめざして—

和泉元千春、上田和子、廣利正代

〔キーワード〕 言語テスト開発理論、TLU 課題、プレースメントテスト、記述、
実践知

〔目次〕

はじめに

1. 調査の背景
2. バックマンの言語テスト開発理論
 - 2.1 言語テスト開発の枠組み
 - 2.2 テストの設計
 - 2.3 <目標言語使用 (TLU) 領域>と<目標言語使用 (TLU) 課題>
 - 2.4 <言語課題の特性>の記述と<構成概念>の定義
 - 2.5 テストの有用性
3. 事例：大学院生研修のプレースメントテスト (インタビュー) 設計の記述
 - 3.1 大学院生日本語研修の背景
 - 3.2 プレースメントテスト (インタビュー) の<テストの目的>
 - 3.3 <TLU 領域>と<TLU 課題>
4. 考察
 - 4.1 <テスト課題の特性>と有用性評価
 - 4.2 テスト結果の利用
 - 4.3 教師の実践知の側面から
 - 4.3.1 実践知を知る過程：教師の経験を記述すること
 - 4.3.2 実践知の暗黙性

おわりに

はじめに

本稿はバックマンの言語テスト開発理論に照らして、国際交流基金関西国際センター（以下、関西センター）での日本語研修におけるテスト実践例を記述し、その妥当性についての調査、検証を試みるものである。テストについて記述する作業の過程で、日常の業務のなかには特に意識されることなく運用してきた教師の実践的な知識の存在を認識し、さらにそれが言語として

外化されにくいものであることがあきらかになった。

まず理論的枠組みについて述べた後、事例研究として理論に照らしテスト実践を記述し考察する。

1. 調査の背景

関西センターでは専門日本語研修、日本語学習奨励研修を中心に長期、短期の各日本語研修プログラムを実施しているが、⁽¹⁾ その設計時において日本語のカリキュラムやシラバスの設計と同様、学習者評価の一つの手段であるテストをいくつか配置している。それらが学習課題ごとの形成的評価として用いられたり、その総括的評価として運用される一方、研修修了時の学習者評価における情報源の一つとなっていることはいうまでもない。多くの研修プログラムを擁する関西センターのような研修機関では、習得段階にあわせた統一的なテストを開発することも一つの選択肢として考えられるが、研修内容が必ずしも同一ではないという事情もあり、テスト開発はそれぞれの研修を担当する日本語教育専門員（以下、専門員）が各研修のもつ専門性、研修期間やその他の諸条件の異なりにしたがって個別に取り組んできた。その様な背景から、いわば研修ごとに個別に〈タテ〉に作られていたテストを、専門員全体によって〈ヨコ〉で検討するということは行われなかった。

一方、関西センターでは、各研修の内容について専門員間で情報を共有していくことを主な目的として、平成11年度から研修事業課内に研究企画推進班が設けられ、一年に数回の研修会を行ってきた（和泉元 他 2002）。このような研修会が行われるようになった背景に、専門員の研修担当形態がセンター設立当初よりやや複雑になり、一つの研修を専属に担当するというよりは、年度ごとに研修担当者の異動があること、あるいは同年度に複数の研修を担当するなど、より多くの研修に関わる可能性を持つようになってきたことが挙げられる。その点からも専門員らは研修会などを通じて、関西センターで実施されているすべての研修について認識しておくことが重要であると考えられる。

また、専門員らが出張や外部の研修活動に参加した場合、その成果を関西センターの研修現場に報告し還元しているが、これも研修会で行われる重要な活動の一つである。2002年3月には、言語テスト開発理論の第一人者であるL.F.バックマン博士を招聘し、日本語教育学会と国際交流基金関西国際センターとの共催で言語テスト開発理論についてのシンポジウム、講演会およびワークショップが開かれたが、その一連の催しに関西センターからも筆者ら数名の専門員が参加し、テスト理論の重要性を強く認識する機会を得た。

そこで、2002年4月の研修会では、バックマン言語テスト開発理論の紹介と、それを手がかりに、関西センターの実践の中から特にテスト開発に焦点を当て、その妥当性の検証を試みることにした。検証の方法は、バックマン理論のキーワードに従って、関西センターで実施しているあるテストを事例とし、その開発過程における様々な情報を収集し記述する方法である。

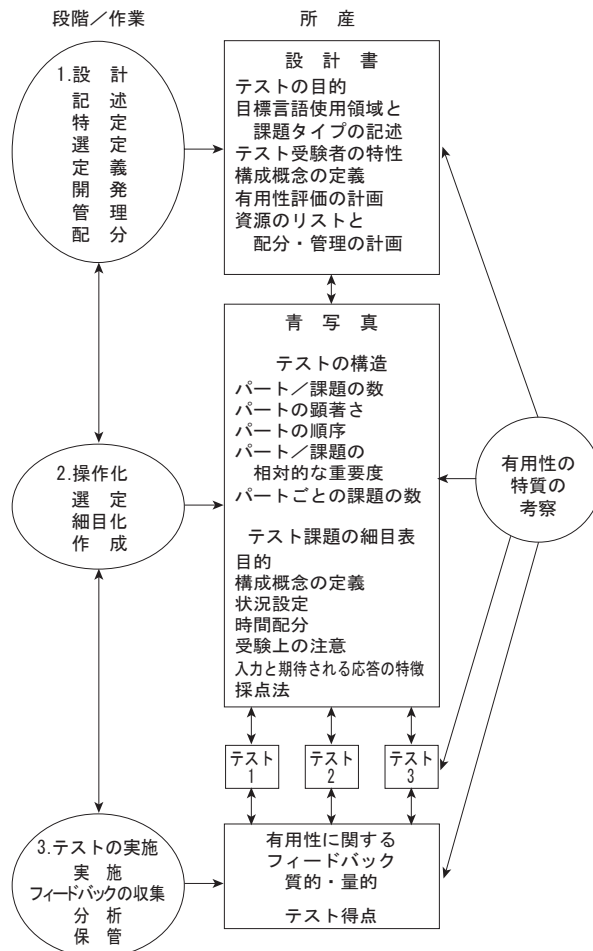
本稿はその研修会で報告された内容に基づき、さらに考察を加えて報告するものである。

2. バックマンの言語テスト開発理論

2.1 言語テスト開発の枠組み

バックマンの言語テスト開発理論によると、それには大きく分けて3つの段階がある。すなわち、<設計（テスト設計の各部門を詳細に記述しテスト結果を有効に利用する確証とする）>、<操作化（具体的なテスト課題を選定し、詳細に記述する）>、<テストの実施>である（図1）。

図1 言語テスト開発の三段階（『<実践>言語テスト作成法』バックマン／パーマー 2000、p.99）



次節では主に<設計>段階に関する理論の枠組みを紹介する。（なお、文中 部分は、バックマンによる言語テスト開発理論の概要部分である。）

2.2 テストの設計

＜設計＞段階はテストの基盤を作る段階であり、まず、＜テストの目的＞を明確にしなければならない。次に、どんな人がテストを受験するのかという＜受験者の特性＞について記述する。

バックマンは、＜受験者の特性＞として

- 1) 個人的特性：年齢、性別や一般的教育水準など
- 2) 話題の知識：受験者の持つ話題の知識
- 3) 言語能力の水準とそのプロフィール：受験者のレディネス
- 4) テスト課題に対する情意反応

の4点を挙げている。

従来、どのテスト開発者も、たとえ詳細な分析までは行わなくとも、それらを念頭に置きつつテストを開発していると考えられるが、さらに＜受験者の特性＞を詳細に記述することによって、受験者の課題遂行にそれらの特性がどのように影響しているかについて、より明確に把握できるようになる。

2.3 ＜目標言語使用（TLU）領域＞と＜目標言語使用（TLU）課題＞

テスト受験者がテストそのもの以外で目標言語を使用する状況を特定し、そこで出会う可能性が高い一連の特定の言語使用課題を記述、列挙する。

バックマンはこれらの課題を＜目標言語使用領域（Target Language Use Domain：以下 TLU 領域）＞での＜目標言語使用課題（以下、TLU 課題）＞と呼び、その記述をテスト開発上で最も重要な作業の一つであると考えている。さらに、TLU 領域は「言語が何のために用いられるか」という観点から、＜実生活領域＞（言語が本質的にコミュニケーションの目的のために用いられる領域）と、＜言語教育領域＞（言語が言語の指導や学習の目的に用いられる領域）とに区別される。

2.4 ＜言語課題の特性＞の記述と＜構成概念＞の定義

＜TLU 課題＞を分析するにあたって、バックマンはそれらの＜TLU 課題＞がどのような特性を持つものであるかを分析し、詳細に記述することを勧めている。具体的には、言語

使用が行われる物理的状況、テスト課題の構造、言語使用者に与えられる入力と言語使用者に期待される応答などに関する特性、および入力と応答との関係が挙げられる。

このように、記述された〈TLU 課題の特性〉によっていくつかのグループに統合された〈TLU 課題〉タイプから、テスト課題としてふさわしいものを選定する。選定されたテスト課題は、〈TLU 課題〉の特性の記述と同様の方法で記述される。

ここで留意すべきことは、〈TLU 課題〉=〈テスト課題〉ではない、ということである。〈TLU 課題〉はあくまで実際の言語使用において見出される課題であり、そのままでは〈テスト課題〉にはならない。また、バックマンが記述を要求している課題特性の中には、テスト開発者が当然の前提としているような特性もあり、これらのすべてを詳細に記述することは一見かなり煩雑な作業のように思える。しかし、一つ一つを改めて仔細に検討することにより、初めて各課題間の弁別特性が明らかになり、課題タイプのグループ分けがしやすくなるのである。そして、この精密な特性記述こそバックマンの言語テスト開発理論の要となるものである。

また、テスト開発者は言語能力のどの側面について測定し、どの側面について測定しないのか、すなわち〈構成概念〉を十分に検討し、定義する必要がある。この段階では〈テストの目的〉に則って、言語能力を構成する要素を抽象的に記述し、定義する。〈構成概念〉に方略的能力や話題の知識を含めるか否かはテストの状況によって異なるため、定義の際には留意しなければならない。

2.5 テストの有用性

言語テストの有用性は以下の 6 つの特質で構成されている。

- 1) 信頼性：測定の一貫性
- 2) 構成概念妥当性：測定したい能力（構成概念）の指標として、あるテスト得点の解釈がどの程度有意義かつ妥当であるか
- 3) 真正性：TLU 課題の特性とテスト課題の特性とがどの程度一致しているか
- 4) 相互性：受験者の課題遂行に、受験者の個々の特性がどの程度どのような形で関わら合うか
- 5) 影響：そのテストが、受験者、教師、社会・教育組織等のテストに関わる人々にどのような影響を与えるか
- 6) 実用性：開発されたテストが物理的に実施可能かどうか

バックマンは開発されたテストの有用性評価にあたって、これら個々の特質の個別評価から評価されるのではなく、それらの特質のバランスに着目して評価されることが肝要であると述べている。その際、どのようなバランスが適切であるかを決定づけるのは〈テストの目的〉である。

以上のような手順を踏むバックマンの言語テスト開発理論では、「いかなる言語テストもある特別の目的、特定のテスト受験者、特定のTLU領域を念頭に置いて開発されるべきである」（前掲書、p.21）と述べられ、〈テストの目的〉、〈受験者の特性〉、〈TLU領域〉、〈TLU課題の特性〉、〈テスト課題の特性〉それぞれを明確に意識することが重要であることが一貫して強調されている。

次節では、関西国際センターで実施されている大学院生日本語研修（以下院生研修）のプレースメントテスト（インタビュー）を、バックマンのキーワードを手がかりに記述し検証する。

3. 事例:大学院生日本語研修におけるプレースメントテスト（インタビュー）設計の記述

3.1 大学院生日本語研修の背景

院生研修は、日本以外の国、地域で大学院修士課程または博士課程に正規生として在籍している者（大半が20歳代半ば～20歳代後半）を対象として、関西センターで実施している4ヶ月の専門日本語研修である。研修参加者は社会科学または人文科学を専攻しており、その研究分野の内容は多岐にわたっている（表1）。それだけではなく、その研究レベルや個々の日本語能力、当研修に参加した目的や期待も多様であるのが大きな特徴の一つだと言える。そのため研修参加者個々のニーズや期待に具体的に踏み込んだコースデザインがしづらいという現状があるが、当研修では大枠の方針を「研究活動の道具としての日本語を伸ばす」という面に焦点を当て、コースデザインを行っている。

本研修では、開始時にプレースメントテストとして、主に文法、読解、作文からなる筆記テストと、聴解能力を測定する聴解テスト、口頭能力を測定するインタビューテストの3部門で構成されたテストを実施し、研修開始時の日本語能力を測定してきた。このプレースメントテストの問題点として、研修開始後に適当でないクラスに配属された研修参加者がいることが明らかになっても、その原因についてテスト設計の側面からは理論的な考察が十分に行えないことが指摘されていた。特に本節で取り上げるインタビュー・テストに関しては、研修参加者1人につき5～6分間、自己紹介として、名前、出身、日本語学習の動機・目的、研究分野についての質疑応答を行っているが、その結果は筆記、聴解テストで測定された日本語レベルと明らかにレベルが異なっていると判断された場合のみ、プレースメントに反映させているにとどまっていた。そこで、これらの課題への解決法を探るため、本調査の対象として、院生研修のプレースメントテストの

表 1 研修参加者の特性と目標言語使用 (TLU) 領域

研修参加者の特性		大学院生日本語研修
A 個人的特性	年齢/性別/国籍	満 22 歳以上 (大半 20 代前半～20 代後半) / 男女 / 多様
	在住状況	長期滞日経験なし。4ヶ月の研修中は宿泊滞在日本以外
	一般的教育水準とタイプ	日本以外の国・地域で大学院修士課程または博士課程に正規生として在籍している者。学術レベルに差はあるが教育水準は概して高いと言える
B 受験者の話題の知識	テストの準備や受験経験の種類と量	研修参加者の中にはインタビューによる口頭能力測定テストに慣れていない者も含まれる
	受験者の話題の知識	専門に関する詳しい知識を持つっていると期待されるが、実際は研修参加者の研究分野、背景 (滞日経験や出身国の学術レベルなど) が異なるので話題の知識にもかなりの違いがある
C 受験者の言語知識の水準とプロファイル	言語能力の全体的水準	原則として初級修了程度～上級
	具体的な文法能力	原則として初級修了程度の文法知識はもっていることが原則。中には能力試験 1 級に合格しているものも含まれる。また能力試験 3 級に合格しているものも口頭能力がきわめて低い者もいる
目目標言語使用 (TLU) 領域の定義	テスト受験に対して予想される情意反応	インタビューによる口頭能力測定テストに慣れない者は心理的抵抗があるかもしれない 来日直後であるため、身体的あるいは心理的疲労から実力を発揮できない場合もあると考えられる
	目目標言語使用 (TLU) 領域の定義	大学院生日本語研修の TLU 課題 (口頭能力) (1) ・日常的で身近な話題について友達と話す (1) ・日常的で身近な話題について日本人研究者と話す (1) ・友達と研究分野について情報交換する (1) ・電話で日常的で身近な話題について友達と話す (1) ・電話で日本人研究者にインタビューの依頼などをする (1) ・研究のため人に会って質問をする (1) ・日本人研究者と直接会って研究分野について話す (1) ・研究分野についてのディスカッションに参加する (2) ・口頭発表のアウトラインについて読みだりしたことを要約する (3) ・初めて会う人に簡単な自己紹介をする (3) ・友達に自己紹介をする (3) ・日本人研究者に自己紹介をする (3) ・研究分野について友達に紹介する (3) ・研究分野について日本人研究者に紹介する (3) ・学会などで聴衆からの質問に答える (3) ・学会やバーチャルなどであいさつやや司会を行う

うち「インタビュー・テストの設計過程」を検証事例として取り上げることとした。

3.2 プレースメントテスト（インタビュー）の〈テストの目的〉

本インタビューテストは、研修参加者が研修を遂行するために必要な口頭能力を、どの程度持ち合わせているかを測定する目的で実施し、その結果によって以下の3点を決定する。

- 1) 目標およびクラス活動に口頭能力が大きく関わると考えられる科目のクラス編成を行う
- 2) テスト結果から研修参加者の口頭能力の傾向を把握する
- 3) 該当科目の開設クラス数に分けられたグループ内の日本語能力の傾向から、予定されているシラバスの手直しを行う

インタビューテスト結果は、研修参加者が参加する研修内容の量とレベルを決定する目的でのみ利用される。

3.3 〈TLU 領域〉、〈TLU 課題〉から〈テスト課題〉へ

院生研修の研修参加者の〈TLU 領域〉について考えると、まず研究分野の内容について日本語で情報収集し、日本語でそれらの情報を整理し、論文執筆や口頭発表など、日本語でなんらかの情報を発信するという〈実生活領域〉タイプの TLU 領域での日本語使用が挙げられる。ただし研修参加者によっては、ある課題が〈言語教育領域〉での日本語使用となる場合もある。つまり、院生研修参加者の中には、研修以外の実生活では日本語で口頭発表をする機会がないため、研修の最終段階で実施している発表会での口頭発表が具体的な〈TLU 課題〉となる者も少なくないからである。

記述作業にあたって、研修参加者のニーズ調査、過去の研修の蓄積から得た情報を参考に〈TLU 領域〉を切り取り、院生研修の〈TLU 課題〉を取り出したうえ⁽²⁾、それらのうち口頭能力に関わるものだけを表1に列挙した。

続いて、課題特性の枠組みに基づき、それぞれの〈TLU 課題の特性〉を記述していくと、「話題」（日常的な話題／研究分野の話題）、「場面」（対面会話／電話会話、相互交渉のある会話／口頭発表）、「言語使用域」（待遇表現など）、という3点から、各課題間の類似点と相違点により、さらに詳細に分類することが可能となった。例えば、院生にとっての〈TLU 課題〉を考える場合、研究活動協力者である大学教授ら研究者とだけでなく、同年代の研究仲間や研究の情報提供者らとの関係構築や維持の上での日本語使用が重要な位置を占めている。そこから、研究分野の内容だけではなく、自己紹介や自国紹介をはじめ日常的で身近な話題について話すことも重要な課題となり得ることがわかる。その結果、院生らはそれらの相手に応じて適切な言語使用域を用いる技能が期待され、〈テスト課題〉として待遇表現を組み込んだものを実施する意味と方法が明らかになってくる。

以上のような記述や分析を含む作業を通じて、〈TLU 課題〉から〈テスト課題〉にふさわしいものを選定した結果、〈テスト課題①〉として「日常的で身近な話題に関する質問に答える」

言語テスト開発過程の記述と検証

と、〈テスト課題②〉として「研究分野に関する質問に答える」という2つの課題を設定できる。

このようにバックマン理論に照らして、〈TLU 課題〉とそこから選定された〈テスト課題〉の特性を記述したのが(表2)である。

実施に際しては、インタビューテストを研修参加者1人につき15～20分程度で行い、試験官(1名)による質問に対する研修参加者の応答を、評価者(1名)が構成概念に対応した項目ごとに5段階評価で採点することとした。また試験官の疲労と集中力の欠如が信頼性に不利な影響を与えないよう、テスト担当者は試験官役と評価者役を交互に担当することとした。

次節ではこのインタビューテストの有用性評価を行い、従来開発運用してきたインタビューテストについて考察する。

表2 目標言語使用(TLU)課題特性とテスト課題特性

		TLU 課題① 日常的で身近な話題について友達/専門家と話す	テスト課題① 日常的で身近な話題に関する質問に答える
言語能力の定義 (構成概念の定義)		①コミュニケーションを妨げない発音の正確さ、流暢さ②文法の正確さ③日常会話に必要な一般的な語彙力④理解力⑤談話構成能力⑥説明能力⑦会話発展能力 < 社会言語学的能力 >⑧形式ばらない言語形式/あらたまった言語形式	①発音—正確さ・流暢さ②文法の正確さ・文型の豊富さ③語彙の正確さ・語彙の豊富さ④理解力(試験官の発話を理解したか?)⑤談話構成能力(単語→単文→複文→連文→段落→複段落)⑥説明能力(わかりやすく論理的に説明できたか?)⑦会話発展能力(積極的に話題を提示したり、膨らませたりできたか?)
課題の特性	物理的条件	問わない	教室
	参加者	言語使用者と友達(あるいは目上の人)	研修参加者1名、試験官1名、評価者1名
	課題の時間	問わない	15～20分(テスト課題①②をあわせて)
	入力特性	音声による。場合によって写真や非言語行動など視覚からの入力もあり	試験官の音声による質問・決められた質問項目+研修参加者からの応答に関する質問基本的な文型によるyes-no質問+基本的な疑問詞を使った質問→背景説明を求める質問一般的な語彙
	期待される応答の特性	音声による。場合によって非言語行動などによる	試験官の質問に口頭で答える(自己紹介、日課、自国紹介)
	入力と応答の関係	言語使用者の応答に応じて対話者は発話の調整を行う対話者の入力によって言語使用者の応答が制限される可能性がある	研修参加者の応答に応じて試験官は発話の調整を行うが、決められた質問は必ず行う試験官の入力によって研修参加者の応答が制限される可能性がある
	話題の特性	日常的で身近な話題	自己紹介 日課 自国紹介
		TLU 課題② 研究分野について友達/専門家に紹介する	テスト課題② 研究分野に関する質問に答える
言語能力の定義		①発音—正確さ・流暢さ ②文法の正確さ ③一般のおよび研究分野の語彙の豊富さ、正確さ④理解力⑤談話構成能力 ⑥説明能力 ⑦会話発展能力 ⑧話題に関する知識(研究分野に関する知識の質と量)⑨形式ばらない言語形式あるいは敬語などあらたまった言語形式	①発音—正確さ・流暢さ ②文法の正確さ ③一般のおよび研究分野の語彙の豊富さ・正確さ④理解力⑤談話構成能力 ⑥説明能力 ⑦会話発展能力 ⑧研究分野に関する知識量(と質)
課題の特性	物理的条件	問わない	教室
	参加者	言語使用者と友達(または日本人研究者)	研修参加者1名、試験官1名、評価者1名
	課題の時間	問わない	15～20分(テスト課題①②をあわせて)
	入力特性	音声による場合によって写真や非言語行動などからの入力もあり	試験官の音声による質問、決められた質問項目
	期待される応答の特性	音声による場合によって非言語行動などによる	試験官の質問に口頭で答える(研究業績と予定)
	入力と応答の関係	言語使用者の応答に応じて対話者は発話の調整を行う対話者の入力によって言語使用者の応答が制限される可能性がある	研修参加者の応答に応じて試験官は発話の調整を行うが、決められた質問は必ず行う試験官の入力によって研修参加者の応答が制限される可能性がある
	話題の特性	研究分野に関する話題	研究分野に関する話題

4. 考察

4.1 <テスト課題の特性>と有用性評価

2.5にあるテストの<有用性>の6つの特質から省察すると、従来行われていたプレースメントテストはある程度その目的を達成しているテストであったことが認められた。

しかし、学習者の状況やテストの目的を<TLU領域>や<TLU課題>、<テスト課題>という鍵となる概念にしたがい、系統だてて把握することで、テスト開発をより具体的に進めることができ、従来以上に幅広い対応が可能であること、また必要であることが明らかになってきた。さらに<有用性>という、より上位の概念から妥当性を検討し、項目ごとの洗い直しによる補強で、従来のテストがさらに質の高いテストとして発展できる可能性があることが視野の中に入ってきた。その意味で、バックマン理論によるテスト開発実践の省察は大変有益であったと認められる。

特に以下のア～ウの3点に関しては、従来のテストを補強するための具体的で即時的な示唆が得られた。

ア. 課題の一部変更

従来のインタビューテストは「日本語学習の経験と目的について説明する」「研究分野について説明する」という2つの課題で行われていたが、<TLU課題>に照らして検討した結果、前者を<テスト課題①>「日常的で身近な話題に関する質問に答える」とし、その下位課題として、「自己紹介する」「日課の説明をする」「自国を紹介する」というより分化した課題を設定するよう調整した。これにより、テストの真正性と言語能力の測定の精度を高めることができる。

イ. 話題に関する知識

従来のインタビューテストでは「話題に関する知識」は<構成概念>として考慮されていなかったが、それを項目として加え、課題ごとに検討することとした。例をあげると、<テスト課題①>「日常的で身近な話題に関する質問に答える」では、話題に関する知識の多少によって日本語による応答の量、質がほとんど影響を受けない課題であるため、話題に関する知識は<構成概念>に含めないが、それに対して<テスト課題②>「研究分野に関する質問に答える」では、母語および日本語等での研究分野の内容に関する知識の量の多少が、一部の日本語科目のクラスワークにも影響を与えるため⁽³⁾、この場合は<構成概念>として独立して定義づけたほうがよいと判断したことなどである。

ウ. 入力の特徴

試験官は、簡単なYes-No疑問文から簡単な疑問詞疑問文、描写、説明を要求する疑問詞疑問文へと質問の型を移行させ、次の質問へ移る時は自然さを心がけて、不自然な入力が研修参加者の応答に与える影響を最小限とするよう配慮しなければならない。これを明文化することにより、従来のいわゆる教師間の申し合わせ事項としての認識ではなく、より厳密な事項とし

て意識することが促される。

また、試験官である研修担当者は院生らの多岐にわたる研究分野の内容に関して、専門的な見地からの入力、つまり専門的な課題を出すことは困難であることが多い。そのため、「研究のテーマは何ですか」、「今までどんな研究をしましたか」などのように、より一般的で、試験官の話題に関する知識の量と質によって課題の質が左右されない入力とすることにした。採点に際しても、研究分野に関する応答の質を発話量で推測せざるを得ないため、実用性は高いが構成概念妥当性は低くなる可能性があるということも否めない。これは、テスト結果の利用の際には評価者の限界も十分に考慮しなければならないという一例であろう。

上記3点に加えて、試験を実施する環境の変化など、状況によっては次の2点についてさらに進化させるための展望が得られた。

エ. 社会言語学的知識

本インタビューでは、研修で社会言語学的能力が直接関係するような科目は必修科目ではないこと、またテスト時間に制限があるという実用性の点から、構成概念として社会言語学的能力を取り上げなかった。しかし、〈TLU 課題〉間の相違点の記述という作業を通して、言語使用域の相違、すなわち社会言語学的能力からの弁別を〈テスト課題〉として取り上げることが可能であることがわかった。

形式ばらない言語形式や、敬語などの改まった言語形式を使用する能力は、研究活動に関わる人たちとの人間関係の構築と維持のために不可欠であるという認識に立ち、今後必要があれば、〈テスト課題③〉として「会ったことのない大学教授にインタビューの依頼をする電話をかける」といったロールプレイなどを課すことも可能である。

オ. 方略的能力

本インタビューテストでは、テスト時間にも制限があるため、研究分野について説明する能力から二次的に「質問できるかどうか」を測定するにとどまり、その方略的能力を〈構成概念〉に含めないこととした。今後、効果的に質問をするという方略を重要視する場合は、方略的能力を構成概念に含め、課題に「○○について試験官に質問する」等を新たに加える必要があるだろう。

4.2 テスト結果とその有効利用

〈テストの目的〉を〈構成概念〉にどのように反映させていくか、〈テスト課題〉としてどの課題を取り上げるか、すなわち〈TLU 課題〉をいかにテストに反映していくかは、常に問題となる点である。

院生研修の場合、〈テストの目的〉を理解し、研修参加者のニーズをテスト課題に反映させることの重要性については、これまで専門員の教師としての経験から認識されてきた。しかし、従来のインタビューテストでは、研究分野に関する質疑応答がなされていたにも関わら

ず、自らの研究内容に関する研修参加者自身の知識の量と質が、彼らの日本語力、および研修活動に影響していることに関して、試験官、評価者らの間に共通認識が必ずしもあったとは言えない。テストがクラス分けに十分反映されないという問題の一因が、そこにあったと推測できる。これは、テスト開発者と試験官、評価者、研修実施担当者の全員が、同じくテストの目的と構成概念のもとでテストを計画、実施し、またその結果を利用するという姿勢を持たなければ、結果を有効に利用できないことを端的に表している一例であるといえるだろう。

4.3 教師の実践知の側面から

4.3.1 実践知を知る過程：教師の経験を記述すること

本調査では、インタビューテストの設計過程を記述することにより、従来のテストの評価すべき点、補強すべき点が明らかになった。一方、その過程でテスト開発者、すなわち教師の「経験的な知」が存在することも明らかになってきた。記述作業は、テスト結果に関わるすべての者の〈テストの共通認識〉を構築するうえで重要な役割を担っており、それがテスト結果の有効利用を可能にするという示唆も得た。

「記述」は教師個人の経験的な知を、教育現場で共有する知へと展開させる方法の一つであるにとらえ、以下、教師の経験と記述作業について、さらに考察を続ける。

本調査のような教師の仕事の記述することへのもう一つの動機付けに、ドナルド・ショーンが提示した「反省的实践家」という概念がある。「反省的实践家 (Reflective Practitioner)」とは、「思考と活動、理論と実践という二項対立を克服した専門家モデルであり、その知を支える鍵は、「行為の中の知」と「行為の中の省察」、「状況との対話」という三つの概念である (ショーン、2001、秋田、2001) という。それらは、専門家あるいは実践者が実践を経験することそのものによって構成される知であり、実践の中で立ち止まってその行為について考えること、そしてそれらを批判し評価する過程で生まれる違和感を受け止めつつ、なぜそのような事態になるのかという問いに答えを出すために、状況の中で新たな枠組みを模索していく過程の知であるとも言え換えられる⁽⁴⁾。

「実践者が自らのフレーム（枠組み）に気づくようになると、実践の現実にフレームを与える別の方法の可能性にも気づくことができる (前掲書、p. 177)」のであるとすれば、逆に、外側にある別のフレームから自らを照らすこと、すなわち「フレーム分析」により、「実践者が彼らの暗黙のフレームに気づくのを助け (前掲書、p. 180)」ることになると考えられる。これらを本調査に置き換えてみると、実施しているテストをバックマンの言語テスト開発理論という外側のフレームに照らして記述するということは、実践者である専門員がおこなってきた研修に存在するであろうフレーム、すなわち実践的知識の枠組みを明らかにしていく方略となりえると考えられる。

4.3.2 実践知の暗黙性

ショーンはさらに『専門家の知恵』のなかで、実践という言葉の持つ意味を「ある範囲の専門的状況における達成的な行為を指し」、それは「達成への準備」を指すと同時に「繰り返し」という要素を含んでおり、その過程で専門家は熟達していくが、実践が安定的である限り、実践家は次第に驚かなくなっていく、つまり「実践知」はますます暗黙で無意識的になり、その結果、自分が今していることについて考える重要な機会を逃がしているかもしれない、と指摘している（前掲書、102-105）。

教師ら（関西センターの場合、専門員および講師）は、日常的に研修の設計、運営、実施という大きな枠組みの作業から、時間単位の授業運営、さらに研修参加者の学習相談や論文作成指導など、実に様々な領域にわたる業務に当たっている。また業務は固定的ではなく、研修参加者が年々変わると同様、担当者はチームの編成、再編成を繰り返しつつ、時に複数の研修を担当することもある。その繰り返しは教師である専門員らを＜熟達者＞として育ててはいくだろうが、ショーンの言葉に照らせば、研修の経験が積み重なっていくにしたがって、専門員らの実践知は、まさに研修という文脈に隠れ、言語化されていく機会は益々少なくなっていくという危険性ははらむことになる。

関西センターにおいて、不定期ではあるが研修会を持ち、それぞれの研修で行われている内容について情報交換を行ってきたことは、いわば暗黙裡に潜んでしまう実践知を言語化し、共有していこうという試みの一つであったと評価できる。一方、テスト理論を援用して教師の実践を記述、分析するという本調査の試みは、研修現場には教師の経験や実践に基づいたある枠組みが構成されていることを明らかにすると同時に、それはややもすれば個人の経験のうちに留まるおそれがあることも、同時に映し出す作業であったと言えよう。

おわりに

Lynchはプログラム評価において妥当性を高めるための方策として、Denzinが唱えた様々な角度からのデータや切り口によって検証していく技法である＜Triangulation＞を紹介している（Lynch、1996）。その視点に立てば、実践の中から＜テスト＞という一つの事例を選び、記述し検証していくことは、教師による自己評価、あるいはプログラム評価の一つの技法としても位置付けることができるだろう。

本稿でとりあげたインタビューテストは、言語テスト開発理論にしたがって作成されたものではなく、それまでの教師としての経験に基づき開発されてきたものである。教師らはそれぞれの経験から、どのように思考しどのように実践してきたか、そのプロセスの中に教師の暗黙の知や技能があったのではないだろうか。本調査により言語テスト理論からそれらのテストが妥当であったと認められたということは評価に値するものといえよう。しかし、さらに重要な

ことは、教育現場の実践を記述していくことの中には、教室の文脈に潜んでいる教師の実践知を再認識し、さらに外化させるための技法としての意味があるということを確認できたことである。その具体的な内容、構成、技法などを解明するためには、今後より焦点を絞った質的な調査が必要であろう。

ただし、それでもなお潜伏してしまうものが、同じ文脈で働く同僚のあいだには、あるいは同僚であるからこそ存在するのかもしれない。もちろん、実践者が経験しているすべてを言語化し、伝達していくことは不可能である。それは暗黙知という概念を唱えたポラニーのいうように「我々は語ることができるより多くのことを知ることができる」（ポラニー、1980）のが現実であり、本質だからであろう。

本調査で用いたバックマンのテスト理論の場合だけでなく、理論はそれを手がかりに現象を分析し記述することも可能にする。それによって調査者個人だけでなく実践現場を共にする集団としての教師の職業的内省を引き起こす手法にもなりうると考える。しかし、それら経験に基づいた知や思考様式という内容そのものだけでなく、それを教育現場の共有財産として育成していく「方法」も、実は身近な日常に潜んでいるのではないだろうか。その重要性を認識することは教師個人においても、さらに研修運営においても極めて重要なことである。

〔注〕

- (1) 国際国流基金関西国際センターでは2002年度現在、長期研修として「司書日本語研修」「研究者日本語研修」および「大学院生日本語研修」「外交官日本語研修」「公務員日本語研修」があり、また短期研修として「日本語履修大学生訪日研修」「日本語成績優秀者研修」「高校生日本語学習奨励研修」のほか、いくつかの奨励研修を行っている。
- (2) 平成13年度に実施した院生研修では、研修開始時に< TLU 課題 >としてあげた項目を含む「大学院生のための日本語能力」という日本語能力、技能の一覧を開示し、研修開始時、研修中、研修後の日本語学習に関する自己目標設定の参考にさせた。
- (3) 平成13年度大学院生日本語研修では、スピーチ技能、談話構成能力養成のために「発信のための日本語」という科目を必修科目として設定した。高いレベルのクラスでは、主に研究分野に関する話題についてスピーチを行うこととした。
- (4) 教師の実践的な知については、「創造的な熟練教師が①即興的思考、②状況的思考、③多元的思考、④文脈化された思考、⑤枠組みの再構成、という五つの特徴で表現される『実践的思考様式』を生成し機能させている」という指摘がある（佐藤、1997）。また、教師のReflection（内省、反省、省察）の方法としてはジャーナル、アクション・リサーチなど様々な手段が紹介されている（Moon 1999, Bailey 2001）

〔執筆分担〕

和泉元千春：3

上田和子：はじめに、1、4

廣利正代：2

〔参考文献〕

秋田喜代美（2001）「解説 ショーンの歩み - 専門家の知の認識論的展開」、『専門家の知恵』ドナルド・ショーン著、佐藤学、秋田喜代美訳、ゆみる出版 p.211-227

和泉元千春、上田和子、大隅敦子（2002）「国際交流基金関西国際センター平成12年度研究企画推進班活動報告」、『日本語国際センター紀要』第12号、国際交流基金日本語国際センター p.123 - 130

佐藤学（1997）「実践的探求としての教育学」、『教師というアポリアー—反省的实践へ—』世織書房、p.135-156

ショーン、D. 著、佐藤学、秋田喜代美訳（2001）『専門家の知恵—反省的实践家は行為しながら考える—』ゆみる出版

バックマン、L.F./パーマー、A.S. 著、大友賢二/ランドルフ・スラッシャー監訳（2000）『＜実践＞言語テスト作成法』大修館書店

ポラニー、M.、（1980）『暗黙知の次元』紀伊国屋書店

Bachman, L.F.,/ Palmer, A.S., (1996) Language Testing in Practice : Designing and Developing Useful Language Tests, Oxford University Press.

Bailey, A.,/ Curtis, A., /Nunan, D. (2001). 2 Self-Awareness and Self-Observation: Cornerstones, 3 Reflective Teaching: Looking Closely. Pursuing Professional Development: The Self as Source, Heinle & Heinle p22-47.

Moon, J, (1999). 6 Reflective Practice in the professions- a practical stance. Reflection in Learning & Professional Development. Kogan page Ltd. UK, p66-78.

Lynch, B,(1996). 3 Validity. Language Program Evaluation. Cambridge University Press. p41-69

