

# 日本語研修生の試験結果から見たコース評価

三浦 香苗・古本 裕子

## はじめに

金沢大学日本語研修コースでは毎週水曜日に定期試験（計14回）、コースの最後に修了試験を実施している。また、平成9年3月には修了生を集めて追跡調査をした。試験は学生の学習を促すだけでなく、コースの見直しや指導の見直しの良い材料となる（日本語教育学会 1991 a）。本稿では次のような視点から試験結果を分析し、最後にコース評価を試みる。

- (1) 定期試験と修了試験の結果は関係があるか。
- (2) 第2期生と第3期生では成績に差があったか。
- (3) 筆記、聴解、口頭試験結果はどのように関係しているか。
- (4) コース修了後の学生の日本語能力はどのように変化するか。

## I. 定期試験と修了試験（日本語能力試験3級）の結果の関係

最初に定期試験と修了試験の概要を示し、二つの試験結果の関係について分析する。

### 1 定期試験と修了試験の概要

定期試験は、既に学習した課の中から範囲を区切って学生に予告し、その範囲の重点項目を中心にした問題を担当教師が作った。定期試験の筆記は、助詞や動詞の活用などを問う文法的な問題、教科書の会話文の中にある「呼びかけ」「感謝」「ことわり」「言いよどみ」などの機能を持つことばを書かせる問題、長文読解問題、作文が主なものである。聴解試験は、文法的な項目を聞き取る問題、未習のことばも含むダイアログやモノログから必要な情報を聞き取る問題が出された。口頭試験はその範囲内で学習した機能を使うロールプレイやタスクを課し、一人の教師が文法の正確さ、語彙の豊富さ、発音、談話構成などを点数化して評価を行った。

一方修了試験としては日本語能力試験3級の問題（平成6年度版）を実際と同じ要領で行った。この試験は国際交流基金が毎年1回行っており、問題と結果が公開されている（国際交流基金 1995）。このテストで60%以上を取ると3級合格が認定されるが、こ

の認定は一般的に権威あるものとされている。日本語能力試験3級の認定基準は、基本的な文法・漢字(300字程度)・語彙(1500語程度)を習得し、日常生活に役立つ会話ができ、簡単な文章が読み書きできる能力(日本語を300時間程度学習し、初級日本語を修了したレベル)である。

日本語研修コースでは、日本語を全くのゼロから始めた学生に対し『新日本語の基礎 I・II』・『Basic Kanji Book No.1』を使って初級レベルを教授している。日本語能力試験で出題される文法項目・語彙・漢字は、コースで提示される項目を越えている。なおかつ、これは彼らがコース修了後に要求される「大学院で研究ができるレベル」には遠く及ばないが、彼らにとっては3級レベルは現実的な目標の一つである。

## 2 定期試験と修了試験(日本語能力試験3級)の違い

この二つの試験の大きな違いは、前者が『新日本語の基礎 I・II』を基準にして、学習した内容が身についているかを見るいわゆるアチーブメントテスト(achievement test)であり、後者は範囲を指定しないプロフィーシェンシーテスト(proficiency test)として使われた点である<sup>1)</sup>。また、日本語能力試験は全て記号選択の問題で、口頭試験がないなどの特徴を持っている。

二つの試験の違いをまとめると表1のようになる。

表1 定期試験と修了試験の違い

	定期試験	修了試験 (日本語能力検定試験3級)
実施	毎週計14回	1回
作成	担当教師	国際交流基金
測定スキル	筆記・聴解・口頭	文法読解・文字語彙・聴解
記入方法	選択式・記述式	選択式
満点	100点	400点
時間	90分	140分
受験者数	第2期生9人・第3期生9人	16269人(平成6年度)
テストの種類	achievement test	proficiency test
範囲	『新日本語の基礎 I・II』の相当課	初級を修了したレベル

## 3 定期試験と修了試験の結果

以下に二つの試験を受験した金沢大学留学生センター日本語研修生第2期、第3期の学生の試験結果の平均値を示す。なお、第2期と第3期は定期試験の一部に変更を加え

たが、基本的には同じ問題である。

表2 定期試験14回分 学生9人の平均

	筆記	聴解	口頭	合計
第2期 (9人)	30.9	22.8	25.1	78.8
SD	6.6	3.9	3.0	13.1
第3期 (9人)	31.7	23.4	25.0	80.1
SD	6.9	4.2	3.2	13.8
満点	40	30	30	100

表3 修了試験(日本語能力試験3級)の平均

	文法・ 読解	文字・ 語彙	聴解	合計	合格
第2期 (9人)	128.3	58.1	58.3	244.8	5人
SD	19.8	15.5	13.7	45.2	
第3期 (8人)	135.1	62.0	57.6	254.8	7人
SD	32.1	19.1	16.4	64.4	
満点	200	100	100	400	

#### 4 定期試験と修了試験の結果の相関

定期試験と修了試験で、測定されている対象者が同じであればその結果には相関があるかもしれない。日本語能力検定試験とその1週間前に行われた第14回定期試験(範囲『新日本語の基礎Ⅱ』45課から50課)を受験した学生8人(第2期・第3期とも)の成績は、試験の性質の違いにも関わらず、高い相関関係(第2期生  $r=0.773$ 、第3期生  $r=0.883$ )を示した。

定期試験と、日本語能力検定試験が高い相関を示すことは、14回定期試験を見れば検定試験の結果をある程度予測できるということを意味する。しかしながら、検定試験問題を受験することは次のような意義がある。

第1には、日本語能力検定試験は一般的な基準として定着しているため、同じ試験問題を同じ条件で行って得た得点は、対外的にも客観的な基準としやすい点である。

第2には、定期試験は実際の実力差よりも少ない値でしか点数に反映しないようになっていることである。例えば、第3期の学生A、Bの定期試験と、日本語能力検定試験の結果を下表4に示す。学生AもBもまじめによく勉強し、定期試験ではいつも満点に近い点を取っているためここでは両者の差は表れない<sup>2)</sup>。しかし、検定試験では学生Bのほうがかなり得点が高く、差が出る。学生Aは日本語未習で来日、普段の授業の観察では習った以上のことはできない。学生Bは『Situational Functional Japanese Vol. 1』を終えて来日して、研修コースでは上クラスで勉強しており、定期試験に解答する力以上の日本語能力を持っている。そのことが表4からよくわかり、これは教師の実感とも一致するものである。

表4 学生A、Bの定期試験と日本語能力検定試験の結果

	14回定期試験合計点	日本語能力検定試験3級
学生A	98点	273点
学生B	93点	345点

一方、かなり努力しても成績が振るわない学生については、定期試験では少しぐらいのミスは減点しないという教育的配慮をしているが、日本語能力検定試験は記号選択式なのでそういった配慮はいっさいできない。そのため、客観的という観点からはいい資料を得やすいことになる。

定期試験と修了試験の結果は強い相関があるが、測っている能力は違う。学生の総合的な日本語能力を判定するには、定期試験のほうが優れているが、客観性という観点から見ると修了試験（日本語能力試験3級）のほうが優れている。よって、学生の日本語能力を判定するにはこの両方の試験結果を材料として使うほうがいいようである。

## II. 第2期生と第3期生の試験結果の差

次に、第2期生と第3期生に実力差があったのか、それは試験の成績に表れていたのかどうかを調べる。

### 1 第2期生と第3期生全体の試験結果の差

表2で定期試験の結果を示したが、第2期生と第3期生の平均点には有意な差はなかった。 $(P > 0.05)$ 。表3の日本語能力検定試験の平均点でも有意な差は見られない $(P > 0.05)$ 。

しかし、第2期と第3期で違う点は、検定試験で合格点に達した人の割合である。第2期では9人中5人(55.6%)が3級に合格できる結果だったが、第3期では8人中7人(87.5%)が合格点で第3期生が第2期生を上回っている。ただし、これも $\chi^2$ 検定の結果は有意に達していない。 $(\chi^2=2.08, P=0.149)$ 。このことから、第2期生と第3期生の成績は全体としては顕著な差がなかったように見える。

### 2 日本語未習者の成績の伸びについて

ただし、上記の結果は第2期生、第3期生が修了時に身につけていた日本語能力の一部を測ったもので、来日した時点で学習者が持っていた日本語能力について考慮してい

ない。よってこの数字はどちらのコースがよかったかを判定する目安にはできない。なぜならば既習者が多ければ終了時点の成績も当然ながらよくなるはずだからである。

この点を考慮するために、第2期と第3期の学生のうち来日時に日本語が未習であった学生のみ（第2期生が5人、第3期生が4人）の修了時の成績を比べると表5のようになる。

表5 日本語未習者の成績の伸び

	未習者数	第14回定期試験		修了試験(検定試験3級)	
		平均	SD	平均	SD
第2期生	5人	72.1	14.3	228.6	38.0
第3期生	4人	84.2	5.3	272.5	6.2
第2期生と第3期生の差の検定結果 (Unpaired t-test)		t=1.59, df=7, P=0.16		t=2.26, df=7, P=0.06	

これを見ると、修了試験においては第3期生の未習者の平均点のほうが第2期生の未習者の平均点よりも伸びる傾向があったことがわかる (P=0.06)。

日本語を同じ期間だけ勉強したにも関わらず、第3期生のほうが成績が伸びる傾向があった原因は次のようなことが考えられるが、本当のところは不明である。

- (1) 2期生は英語が堪能な者が多く、日本語を使う場面が教室内に限られていた。これに対し3期生は学生同士の共通語が日本語で、授業外でも日本語をよく使用していた。
- (2) 3期生は教員研修生が多く、2期生に比べ、日本語の勉強に対する意識が高かった。
- (3) 3期には、コースの運営が軌道に乗ってきた。

しかし、以上の結果は限られた学生数からみたものである。もしかすると学習者の母語、学習方法、年齢、教授法、クラスの雰囲気などによって成績の伸びに違いが出ているのかもしれない。今後資料を蓄積することによってこれらの影響について検討していく必要があるだろう。

### Ⅲ. 筆記・聴解・口頭試験結果の相関

次に研修コースの定期試験で測定している筆記・聴解・口頭試験の結果がそれぞれどのように関係しているのかを検討する。

筆記・聴解・口頭の3つのスキルが相互に関係しているのは当然のことである。文法

や語彙の力がなくては、筆記だけでなく聴解や口頭試験でも高得点はとれないだろう。また聞き取りの力がなくては会話はできない。しかし、一方では3つの試験は違う能力を計っているのも確かである。筆記試験ができたからと言って会話がうまくできるとはかぎらないのである。現在までの研究では、これらがどのように関係しているのかは明らかでない。

ここでは、定期試験の3つのスキルの獲得点数の相関を見てみることにする。以下表6に、第2期生の13回分（受験人数が非常に少ない第13回の試験結果を除く）の筆記と聴解、筆記と口頭、聴解と口頭の点数の相関係数、それと修了試験の筆記と聴解の相関係数の平均を下の表6に示す。

表6 第2期生の定期試験の筆記と聴解、筆記と口頭、聴解と口頭の相関係数

	13回の相関係数の平均	SD
①筆記と聴解 相関	0.671	0.181
②筆記と口頭 相関	0.871	0.054
③聴解と口頭 相関	0.669	0.180

①と②の差 ( $t = 3.82$ ,  $df = 25$ ,  $P < 0.01$ )

②と③の差 ( $t = 16.717$ ,  $df = 24$ ,  $P < 0.01$ )

①と③の差 ( $t = 0.04$ ,  $df = 25$ ,  $P > 0.1$ )

ここで調べた三つの相関はいずれも高いが、特に②の筆記と口頭試験の相関が非常に高い。また、②の相関と他の二つの相関の平均値の間には有意差があった（Unpaired t-test）。

筆記と口頭の相関が特に高いのはなぜであろうか。その理由として次のようなことが考えられる。

- (1) 実際に筆記能力と口頭能力に普遍的に相関がある。
- (2) この試験がアチーブメントテストであるため、筆記試験と口頭試験で問われる問題が共通している場合がある。
- (3) 口頭試験の評価をする場合、筆記試験と同じような項目ができた場合によくできたという判断を下しやすい。
- (4) 聴解試験はコミュニケーション的な問題が多く、他の試験と性質が少し違い相関が低い。

試験の結果は問題の出し方、評価の方法によって大きく左右されるものなので、(1)のような結論を出すまえに、テスト問題の改良、評価基準の見直し、多人数で評価することによる客観化などが是非とも必要であろう。

第4期(1997年4月～9月)では、上記の(2)(3)に注目する。口頭試験の評価に際して、文法事項を正確に言えたかどうかだけでなく、談話の運用力などに配点を多くして経過を見たい。

#### IV. コース修了後の学生の日本語能力

学生がコース修了時に身につけていた日本語の能力は、コース修了後どのように伸びていくのだろうか。3月4日に日本語能力の追跡調査を行なった。そこでは、コース修了後の日本語の学習状況・生活状況などインタビュー調査した。

このうち第2期生3人が自分の日本語力を知るために筆記テストをすることを希望した。学生3人はいずれも修了時に日本語能力試験3級合格の基準に辛うじて達していたが、追跡調査の時点で2級を受験できる力はないと判断した。そこで、前回受験したときとは違う年度(平成5年度版)の日本語能力検定試験3級を受験させた(国際交流基金 1994)。二つの年度の問題には難易度に多少違いがあったため、両年度の学生の素点を偏差値に直した数値を下に示す。

表7 コース修了時と6カ月後の日本語能力検定試験の結果

	読解・文法		聴解		文字・語彙		合計	
	修了時	6カ月後	修了時	6カ月後	修了時	6カ月後	修了時	6カ月後
学生A	48.45	58.86	54.44	62.20	37.92	59.89	47.76	61.64
学生B	50.20	54.76	53.04	62.20	48.21	56.37	50.75	58.47
学生C	50.70	41.59	59.57	62.20	37.24	45.80	54.42	48.45
平均	49.78	51.74	55.68	62.20	41.12	54.02	50.98	56.19
SD	1.18	9.02	3.44	0.00	6.15	7.33	3.34	6.89

文字・語彙と聴解については全員が成績がよくなっている。特に、聴解は3人の平均値も6カ月後によくなる傾向があった(Paired t-test,  $P=0.083$ )。ところが文法については学生Cの成績は急激に下降しているのである。

このような差が表れるのは、コース修了後の学生の生活に起因するものであろう。インタビュー調査からわかったことであるが、学生AはBコース修了後も継続して日本語の授業に出席したり、積極的に日本語を使用したりしている。一方学生Cは日本語の授業には出席しているものの、日本人と会話すること自体が少なく、専門ではほとんど日本語を使用しないという状態であった。

文法力に関しては、普段の会話はさほど文法力がなくても成立するため、日本で生活していても意識的に勉強しない限り忘れていく一方なのではないだろうか。これに対し、聴解力や語彙力などは特別に勉強しなくても日本に住んでいれば自然に力がついてくる種類のものである。このことが6カ月後の差に表れていると推定できた。

村上（1996）は名古屋大学の研修生を追跡調査し、日本語の使用の困難度をアンケート調査した。その結果、来日2年以上の学習者は「話す」「聞く」ことに対してはほとんど問題がなくなるが、「読む」「書く」については2年未満の学生と困難の度合いにあまり変化がないことを確かめている。今回の調査からみると、その原因が主として文法力の低下である可能性が伺えた。

測定できた人数が少なく、この結果を直ちに一般化することはできないが、この調査によってコース修了時に身につけていた日本語力のある面は、そのまま順調に伸びていくばかりではないことが分かった。今後は学生が実際に必要としている日本語力とは何かなどについても調査・研究が必要であろう。

## V. 試験結果によるコース評価

最後に、試験で学生が獲得した点数は学生が身につけた日本語能力とみなし、これによって我々のコースの評価を試みる。

ここで注目するのは、日本語能力検定試験3級で合格点に達した人の数である。受験した17人中12人が合格点に達している。6か月間（実質17週間）のコースで、ゼロから始めて初級の課程を満足できる成績で修了することは、そう容易なことではないだろう。その上、このコースではプレゼンテーション・プロジェクトワークを学生に課しており、かなりの時間数と労力を費やしている。これも学生の日本語能力を向上させるためのものではあるが、即効性はなく、ましてや文法力を重視する日本語能力検定試験対策となるものでもない。それにも関わらず過半数の者が日本語能力検定試験3級に合格する実力を示したということは、一応基準を満たしたと評価してもよいのではないかと考える。

## VI. まとめ

以上、本稿では日本語研修コースの学生に実施した試験の結果を分析した。はじめに、定期試験と修了試験（日本語能力試験3級）に現れる日本語力の違いについて分析したが、これらの試験は互いに相関を持ちながらも違う種類の力を測定する試験で、双方とも受験する意義があることがわかった。

次に第2期生と第3期生の日本語能力の差について分析した。両者には全体のグループ差は見られないが、第3期生には3級合格者が多いこと、ゼロから日本語を勉強した学生の伸びが第3期生の方が大きい傾向があり、第3期のほうがコースとして成功した可能性があることを述べた。

口頭試験結果については筆記試験結果と高い相関があることがわかったが、その評価の基準などをどう扱うかなど問題が山積しており(日本語教育学会 1991c, 庄司 1995)、今回は詳しい分析ができなかった。

最後にコース修了後の学生の日本語力の伸びは、聴解力が伸びる傾向にあったが、文法力については下がっている学生がいることがわかった。日本語力は修了後の日本語使用の仕方に大きく依存すると思われた。

試験結果から見ると、コースは一応の基準を満たしていると自己評価した。今後は口頭能力・コース修了後のニーズ分析なども視野に入れた資料を蓄積し、コースの改良に役立てて行きたい。

最後に、検定の方法など有益なアドバイスを頂いた岡澤孝雄氏に深く感謝の意を表したい。

#### 〈注〉

- 1) 二種類のテストの違いについては 日本語教育学会 (1991b, p.103) 参照。
- 2) このような現象が起きたとしてもテストの妥当性を疑う必要はない。なぜならば、このテストがアチーブメントテストで、その範囲内の学習項目の定着したかどうかを測定するのが目的だからである。(日本語教育学会 1991b, p.104)

#### 〈参考文献〉

- 国際交流基金 (1994) 『日本語能力試験の概要 1994年版 (1993年試験結果の分析)』
- 国際交流基金 (1995) 『平成6年度日本語能力試験 分析評価に関する報告書』
- 庄司恵雄 (1995) 『日本語研修コース 標準口頭表現力測定検査法開発のための企画調査 研究成果報告書』平成6年度科学研究費補助 (総合研究 (B)) 研究』岡山大学
- 日本語教育学会 (1991a) 『日本語教育機関におけるコース・デザイン』凡人社
- 日本語教育学会 (1991b) 「第2章 よいテストとはどのような性質を持つか」『日本語テストハンドブック』大修館書店 10-130.
- 日本語教育学会 (1991c) 「第9章 対面テストと作文の出題と採点」『日本語テストハンドブック』大修館書店 292-305.
- 村上京子 (1995) 「日本語研修コースにおける成績と修了後の困難度の評価」『日本語研修コース修了生 追跡調査報告書2』名古屋大学留学生センター 研究代表 尾崎明人 文部省科学研究費基盤研究 (B) 6-13.

## **Evaluating the Intensive Japanese Course for Beginners Using the Test Results**

Kanae MIURA and Yuko FURUMOTO

**ABSTRACT** This paper analyzes the results of the weekly achievement tests (oral, listening and writing) and the proficiency test (level 3) for the course. Two groups of students from different semesters took the tests. Group B spoke Japanese even after school, while Group A often showed reluctance in using Japanese. The outcomes are as follows. (1) Difference in scores of Group A and B was not statistically significant. However, Group B was more successful since the number of students who passed the proficiency test was larger, and a tendency of greater development was seen in Group B when zero-beginners were compared. (2) The scores of achievement tests showed close correlation to those of the proficiency test. In spite of this close correlation, these two types of tests are necessary because they detect different aspects of Japanese language ability. (3) The scores of oral tests were correlative to those of written tests. (4) The results of the two proficiency tests given with a 6-month interval indicated that an individual's use of Japanese after the course determined the degree of improvement in grammar, while listening skill improved regardless of whether he/she used Japanese.