東京外国語大学
留学生日本語教育センター論集　22：33～47，1996

学部留学生の日本語能力試験開発のための基礎研究（２）

横田　淳子・伊東　祐郎

（1995．10．31受）

１　研究の意義と目的

学部入学予定留学生用の日本語能力試験を開発するためには、学部留学生に必要な基礎日本語能力を明らかにしていかなければならない。その方法の一つとして、昨年度は、文部省学部留学生の予備教育機関として２０余年の実績を持つ本センターで２学期終了時に学部留学生に対して行っている学内試験の分析を行った。今年度は、外国政府派遣留学生に対する予備教育の一つであるマレーシア政府派遣学部留学生現地予備教育(1)の最後に行われる修了試験を分析の対象とし、学部入学予定留学生の日本語能力試験開発のための基礎資料を作成することとした。

２　研究方法

２－１　分析資料

マレーシア政府派遣学部留学生現地予備教育修了試験は、マレーシア政府によって選抜された高校卒業生が、日本留学のために現地マラヤ大学予備教育部で２年間の予備教育を受けた後、予備教育の修了試験として受けるものである。修了試験は全部で８科目あり、日本語はそのうちの一つである。これらの試験は、予備教育修了時点で日本の大学に入学し得る能力があるか否かを判定する目的で作成されている。

今回、分析の対象としたのは1994年１月に実施された日本語の試験である。受験者数は139名であった。日本語の試験は、聴解、文字・語彙、文法・読解の３種類に分かれ、３種類の試験の合計は200点になっているが、試験結果の分析にあたっては、問題項目による重みづけをせず、各問をすべて１点として処理した。また、文法と読解はひとつの試験のなかに入っているが、それぞれ別個に処理することとし、以下の４種類の試験をそれぞれ分析した。

１　聴解　　２５項目

２　文字・語彙

①漢字の読みを書かせる問題　　４０項目

②漢字を書かせる問題　　４０項目
③語彙を選択する問題　　２０項目
３　文法　　２０項目
４　読解　　３５項目

２－２　分析プログラム

昨年度と同じ国際基督教大学教養学部教育学科石本菅生教授作成のテスト分析用プログラムパッケージを使用し、以下の手順で行った。

① 学生のＩＤ番号940001～940139を作成する。
② 各試験について正答を入力し、正答データを作成する。
書きと読みを記入する文字試験については、正答をａとして入力する。
③ 各受験者の答案結果をコンピュータに入力する。
文字試験については、正答をａ，誤答をｂとする。
④ 入力結果をプリントアウトし、各受験者の答案用紙と照合する。
⑤ 分析プログラムを使って、各試験につき以下のデータを得る。
・個人得点の採点
・統計量の計算（平均、標準偏差、最高点、最低点、得点分布）
・各項目の（成績上位群、成績下位群、全体別）通過率と弁別指数
・各項目の選択肢別回答数（率）
・成績グループ別各項目選択肢別回答数（率）

２－３　項目分析

個々の試験項目を内容面と統計的数量の面から検討した。

内容面の分析のためには『日本語能力試験出題基準』[2]（以下、『出題基準』）を参照し、それぞれの試験項目が日本語教育においてどの程度のレベルに属する学習事項であるかを検討した。文法については『出題基準』の［文法的な＜機能語＞の類－１級（サンプル）］、［文法的な＜機能語＞の類－２級（サンプル）］、３、４級の［文法事項］および［表現意図等］、［１、２級語彙表］、［３、４級語彙表］を使い、文字・語彙に関しては［１級漢字表］、［２級漢字表］および［１、２級語彙表］を使った。

統計的分析のためには、『日本語能力試験の概要(試験結果の分析)』[3]で用いられている基準を参考に項目を以下のように４つの水準に分類した。

【水準Ａ】・正答率が25％未満の項目。
【水準Ｂ】・正答率が25％以上、80％未満で、弁別指数が0.30未満の項目。
【水準Ｃ】・正答率が80％以上で、弁別指数が0.30未満の項目。
【水準Ｄ】・上記の水準Ａ、Ｂ、Ｃに抵触しない、統計的観点からはよい項目。

3　試験結果の分析

3－1　聴解（25項目）

A　試験の内容

聴解試験は異なる形式による３つの問題群（問題Ⅰ～Ⅲ)から構成されている。問題Ⅰは９項目、問題Ⅱは６項目、そして問題Ⅲは10項目、全25項目である。

問題Ⅰは、絵やグラフが提示されていて、それらの視覚情報から正しいものを選ばせる問題である。形式は最初に“Pre-Question”が音声で与えられ、その後に絵やグラフに関する内容文が続く。その後、同じ質問が“Post-Question”として再度与えられ、四肢択一で答えるものである。

問題Ⅱは、視覚情報は与えられず、音声のみを聴いて答える問題である。各項目ともはじめに音声で“Pre-question”が与えられ、その後に会話などの内容文が続く。同じ質問が再度“Post-question”として与えられた後、４つの選択肢が音声で与えられ、正しいものを選ばせる形式のものである。

問題Ⅲは２問あり、それぞれ５項目から構成されている。各問ともまず最初にテープで比較的長い内容のものを２回聴かせ、その後、５つの短い音声情報が与えられ、情報の内容についての適否を○×で答える形式のものである。“Pre-question”が与えられないので、受験者は全く予備情報がない状態で、テープ音声を聞くことになるが、メモをとってもかまわないことになっている。

B　全体像

聴解試験の結果の概略は以下の通りである。

有 効 数：136
総項目数：25
平　　均：17.28
標準偏差：3.33
最 高 点：23
最 低 点：9

試験全体の分析結果を見ると、平均が25点満点で17.28点（69.10%）と約７割である。総点の得点分布から受験者集団が高得点側に寄っていることがわかり、試験問題が比較的易しかったことがわかる。全体的には試験の難易度に関しては大きな問題はなかったものと考えられる。

C　項目分析

問題Ⅰから問題Ⅲまでを水準で分類してみると表１のような結果になった。

問　題	問題数	水 準 A	水 準 B	水 準 C	水 準 D
問題Ⅰ	9	0	3	1	5
問題Ⅱ	6	1	1	1	3
問題Ⅲ	１０	1	3	1	5
	２５	2（8%）	7（28%）	3（12%）	１３（52%）

（表１　聴解問題）

総項目数の50%以上にあたる13項目が水準Ｄに分類され、比較的良問が多かったことがわかる。次に、弁別指数が低い水準Ｂに属するものは７項目、水準Ｃは３項目、水準Ａは２項目となっている。難しすぎるとされる水準Ａに分類された問題項目の形式をみると、ある言葉がどんな意味で使われているかを文脈から推測させる問題や、話題の流れから若干掛け離れている登場人物について問う問題で、かなりむずかしい設問であることがわかる。問題の難易度の点では、視覚情報が与えられている問題Ⅰと視覚情報の与えられていない問題Ⅱの間では、大きな差は見られなかった。また、問題Ⅲは、“Pre-question/Post-question”形式を持つ問題Ⅰ、Ⅱと比較しても正答率の点で差は示されなかった。また、10項目中３項目(30%)が弁別指数が低かったのは、解答形式が○×であったことが影響したのだろうか。解答形式を検討するうえで、考慮すべき点となろう。

３－２　文字・語彙（100項目）

A　試験の内容

文字・語彙試験は３つの問題群（問題Ⅰ～Ⅲ）から構成されている。問題Ⅰ（40項目）は漢字の読み方をひらがなで書かせる問題である。漢字１字を送りがなつきで１項目として取り扱うものと、２字あるいは最大で３字からなる熟語を１項目として扱うものがある。すべての漢字は、短文の中に組み込まれており、漢字そのものが独立した形では出題されていない。したがって、各問題は有意味な

文章で構成されている。

問題Ⅱ（40項目）は、文中のひらがな部分を漢字に書き換えさせる問題である。漢字の再生能力に関する問題であるため、漢字１字が採点対象単位になっている。したがって、漢字熟語も構成漢字ごとに採点される。

問題Ⅲ（20項目）は、漢字語彙に関する問題である。文の中の空欄に四肢から最も適当と思われる語彙を選び入れ、文を完成させる形式である。

読み方と書き方の問題はすべて実際に記入させる形式であるため、正確な読み方と書き方の能力が測定でき、表面的妥当性（face validity)は高いと考えられる。

B　全体像

文字・語彙試験では、問題形式ごとに３つの試験結果にまとめた。概略は以下の通りである。

	問題Ⅰ（漢字読み）	問題Ⅱ（漢字再生）	問題Ⅲ（語彙選択）
有効数	136	136	136
総項目数	40	40	20
平　均	28.96	30.47	11.33
標準偏差	5.20	6.02	2.69
最高点	37	38	17
最低点	8	8	4

（表２　文字・語彙問題）

試験全体の分析結果から各問題の平均を見ると、問題Ⅰが40点満点で28.96点（72.40%)、問題Ⅱが40点満点で30.47点（76.17%)、問題Ⅲが20点満点で11.33点（56.65%）となり、問題間に難易度の差があることが認められる。総点の得点分布を比較してみると、問題ⅠとⅡはほぼ同じような分布を示している。両問題ともに分布が高得点側に偏っており、低得点の分布に広がりが見られる。一方、問題Ⅲは比較的正規分布に近いことが特徴として挙げられる。このことから問題ⅠとⅡは相対的に易しい問題であったことが考えられる。

C　項目分析

問題Ⅰでは、漢字の読み方の能力測定のために出題された漢字を、『出題基準』に基づいて分類し、各級の漢字数と平均正答率を出した。

漢字レベル	漢　字　数	平均正答率
1級	12文字	57.80%
2級.	26文字	79.47%
その他(*)	2文字	63.95%

（表3　文字・語彙問題Ⅰ）

1級レベルの漢字で読み方の正答率が50%以下のものをみてみると、「争い」(50.0%)、「意図」(19.1%)、「錯覚」(8.8%)、「促す」(2.9%)、「把握する」(0%) が挙げられる。特に20%以下の漢字は、一般の日常会話で使われる語彙と言うよりは、新聞、学術誌等の論説文や、小説などで用いられるかなり高度な語彙とみることができよう。2級レベルの漢字で正答率が50%以下のものには、「怠る」(39%)、「摩擦」(36%)、「情報網」(33.8%)、「補う」(23.5%)が挙げられる。1級レベルの語彙ほど低い正答率ではないが、正答率を低くしている要因には、訓読みと音読みに関する知識の欠如や、複合語、合成語などの語彙力不足が挙げられよう。

出題基準に含まれないその他の漢字(*)では、極端に正答率の低いものはなかった。参考までに挙げると、「陥る」(46.3%) と「回避」(81.6%) の2文字で、後者に関しては非常に高い正答率であった。正答率の一番低い問題項目は「錯覚に陥る」であった。

問題Ⅱでは、漢字の書き方能力を測定する問題である。問題として出題された漢字語彙を、『出題基準』に基づいて分類してみると、以下のようになった。

漢字レベル	漢　字　数	平均正答率
1級	1文字	84.60%
2級	39文字	75.96%

（表4　文字・語彙問題Ⅱ）

1級、2級ともに正答率は高く、受験者の漢字の再生能力の高いことが推測される。正しく書けなかった漢字をみてみると、「承(知)」(0%)、「処(理)」(18.4%)、「(承)知」(19.1%) が挙げられる。「承」「処」「知」は、曲線や「はね」「とめ」を多く含んだ漢字で、部首あるいは図式化パターンの認識が難しい漢字かもしれない。また、正答率0%から推測されることは、対象漢字そのものが全くの未習語であるという可能性とともに、漢字を知っていても熟語そのものが未習のために解答できなかったことも考えられる。

問題Ⅲでは、まず20項目を水準別に分類してみた。結果は以下の通りである。

水　　準	水　準　A	水　準　B	水　準　C	水　準　D
項目数　(割合)	3　(15%)	3　(15%)	5　(25%)	9　(45%)

(表5　文字・語彙問題Ⅲ)

次に、水準別に選択肢で与えられている語彙を『出題基準』と照らし合わせ、実際の問題文に即して、正答率と選択肢の関係を分析する。

水準A：項目(6)では、「学校からの＿＿＿」に「通知」(2級)を選ぶ問題であるが、正しく選んだ者はわずか7.4%で、「通報」(*)を選んだ者が80.1%もいた。これらの語彙の使い分けが難しいと判断できる。また項目(7)では、「試合が＿＿＿した」に「中断」(1級)を選ぶ問題であるが、正答を選んだ者は20.6%で、誤答「中止」(2級)を選んだものが44.9%、「休止」(*)を選んだものが33.1%いた。この問題では、格助詞によっては「中止」も正答になる可能性もあり、読解力と文法力双方が要求される難しい問題となろう。項目(17)では、「銀行が隣町へ＿＿＿した」に「移転」(2級)を選ぶ問題である。正答「移転」を選んだ者はわずか5.9%で、「移動」(2級)を正答として選んだものが75%もいた。被験者の母語では「移転」も「移動」も同じ語彙を使っている可能性があり、語彙の正確な意味特徴と使い方に関する知識が求められるむずかしい問題である。

水準B：項目(2)の「人口が＿＿＿する」という文では、「集中」(2級)を選んで文を完成させる問題である。正答を選んだものが32.4%、誤答「集合」(2級)を選んだものが42.6%いた。項目(3)では、「科学と化学は＿＿＿しやすい」に「混同」(1級)を選ぶ問題である。正答を選んだものは30.9%で、「混雑」(2級)を選んだものが34.6%、「混合」(2級)を選んだものが24.3%いた。項目(12)では、「＿＿＿に寒くなった」に「急激」(2級)を選ぶ問題である。正答を選んだものは57.4%、他の選択肢は11%から19%の幅で選ばれている。困難度には問題はないが、弁別指数の低い問題項目である。

水準C：項目(1)では、「日本の地形の＿＿＿は、～」に「特徴」(2級)を選ぶ問題である。正答を選んだものは83.8%で、「特殊」(2級)は14%、「特別」(2級)は2.2%、「特異」(*)は0%のものがそれぞれ選んでいる。項目(8)の「隣のステレオ音がうるさいので、＿＿＿を言いに行った」では、正答「文句」を94.1%のものが選んでいる。「不足」(2級)「不正」(2級)「欠点」(2級)を選んだものは、それぞれ2.9%、2.2%、0.7%である。項目(9)は、「明日、＿＿＿がありますか」に「予定」(2級)を入れる問題である。正答を選んだものは95.6%で、「予想」(1級)

「予告」(*)「予知」(*)は、それぞれ2.9%、1.5%、0%のものが選んでいる。項目(15)では、「宇宙には_____の星が存在している」に「無数」(2級)を選ぶ問題である。正答は81.6%のものが選んでおり、誤答「無視」(2級)、「無害」(*)、「無料」(2級)はそれぞれ14.7%、2.9%、0.7%のものが選んでいる。項目(18)は、「中学生に対して、_____に注意した」に「厳重」(2級)を入れる問題である。誤答「厳格」(*)「厳正」(*)「厳命」(*)はそれぞれ1.5%、0.7%、0%のものが選んでいる。

意味的特徴が似通っている漢語系抽象語彙を選択させる問題Ⅲでは、日本語と母語における意味的差異、あるいは転移によって、よくできる上位群の学生が誤答を選ぶ場合もあり、一概に正答率の高低のみで難易度を推測するのは危険かもしれない。

3－3　文法（20項目）

A　試験の内容

問題Ⅰと問題Ⅱに分かれ、各問題とも１０項目から構成され、答えはすべて四肢選択による。

問題Ⅰは、一つの文の中の空欄に適切な語句を入れる問題である。選択肢となる語句は、助詞相当句（例：にとって）、接続助詞（例：と）、とりたて助詞（例：さえ）などである。一つの文の長さは最大５０字ほどである。

問題Ⅱは、一つまたは二つの文が与えられ、その文のうちの下線を引いた部分と内容が最も近いものを選択肢から選ぶ問題である。下線が引かれている部分は、形式名詞（例：ところ）、文末表現（例：かねない）、使役・受け身表現、接続助詞（例：たら）などを含むものである。

B　全体像

この試験の結果の概略は以下の通りである。

有 効 数：136
総項目数：　20
平　　均：　11.66
標準偏差：　2.71
最 高 点：　17
最 低 点：　6

試験全体の分析結果を見ると、まず平均が20点満点で11.66点（58.30%）とほぼ6割であり、また、ヒストグラムからも正規分布に近いことが分かる。全体像に関しては、能力試験として統計的には問題がないと考えられる。

C　項目分析

問題Ｉの各問題項目を４つの水準に分類し、さらに各問題項目の４つの選択肢が『出題基準』のどの級の文法項目、または語彙に属するかを検討した（表６）。選択肢の級の欄の1.2.3の数字は選択肢の語句が属している級を示している。＊印は１～４級になく、１級以上と思われる語句である。語句が２つの級の語彙を含んでいる場合は高い方のみを記した。

問題項目	正答率(%)	弁別指数	水準	選択肢の級			
				正答	誤答①	誤答②	誤答③
1	96	0.08	C	2	2	2	＊
2	90	0.14	C	2	2	2	2
3	67	0.22	B	2	1	3	3
4	18	0.03	A	1	2	2	2
5	61	0.39	D	＊	1	1	＊
6	56	0.67	D	2	2	2	2
7	42	0.44	D	2	1	2	3
8	44	0.44	D	2	1	2	3
9	53	0.39	D	2	1	2	＊
10	24	0.08	A	2	2	2	＊

（表６　文法問題Ｉ）

項目１と２は水準Ｃで、易しすぎる問題である。両方とも格助詞相当句を挿入する問題であり、ほとんどの受験者が学習している項目であることがわかる。ただし、試験の最初の問題としてはこのぐらい易しい問題でも構わないであろう。

項目３は正当率は適当であるが、弁別指数が低いため、水準Ｂになっている。

項目４、10は、正答率が25%以下である。特に正当率が低い項目４をみると、まず、正答が１級の文法事項であるために正答率が低くなったと考えられる。詳しく検討するために、受験者を総得点によって、得点の高い順に上位27%を最上位グループ（以下、Ｈ）、次の上位23%を上位グループ（以下、ＭＨ）、その次の23%を下位グループ（以下、ＭＬ）、一番下の27%を最下位グループ（以下、Ｌ）

の４つのグループに分け、それぞれのグループがどの選択肢を選んでいるかを分析していく。すると、この問題に関しては、どのグループも誤答である選択肢ｃを選択している者が多い。特定の選択肢に回答が集中した（40％以上）のは、試験を行った機関における教育に受験者を混同させる要因があったとも考えられる。

項目６は統計的な点ではよい試験項目の見本となるような結果が出た。全体の正当率は約60％であるが、Ｈでは正答を選択した者が89％と高く、ＭＨ、ＭＬ、Ｌとなるにしたがって正答率が低くなり、Ｌでは四肢をほぼ同じように選んでいる。選択肢はすべて２級である。

項目８もＤ水準に含まれるが、１級の語彙である選択肢を選択した者は一人もいない。受験生の知らない１級の語彙ではなく２級の語彙を使えば、統計的にもっとよい問題になったと考えられる。

問題Ⅱの各問題項目に関しては、問題文および４つの選択肢が、どの級に属するかを検討した（表７）。問題Ⅰの場合と同様に、一つの語句・文の中にいくつかの級の語句が含まれている場合は、高い方の級のみを記した。また、括弧で記した数字は『出題基準』の表にはないが、括弧内の数字の級に相当すると筆者が判断したものである。

問題項目	正答率（％）	弁別指数	水準	問題／選択肢の級				
				問題	正答	誤答①	誤答②	誤答③
1	79	0.19	B	2	3	3	3	3
2	74	0.36	D	2	3	3	3	3
3	93	0.08	C	3	3	3	3	(3)
4	46	0.31	D	1	3	3	3	3
5	31	0.44	D	(2)	3	(3)	3	3
6	56	0.56	D	1	2	2	2	2
7	44	0.28	B	3	3	3	3	3
8	68	0.36	D	2	2	2	3	2
9	61	0.33	D	1	2	3	2	2
10	65	0.42	D	(2)	3	3	3	3

（表７　文法問題Ⅱ）

問題Ⅱは概ねＤ水準であるが、易しすぎる問題が項目３、弁別指数の低い問題が項目１と項目７である。

項目３は３級のレベルであり、明らかに受験者のレベルより易しい問題だったと考えられる。項目１もかなり易しい問題で、そのために弁別指数が低くなったのであろう。

項目７は弁別指数が低いために、Ｂ水準になっている。受け身、使役、補助動詞としての「あげる」「もらう」など、すべて３級の文法事項であるが、正確な意味を文脈に即して把握するのは受験生にとって難しいのであろう。Ｈでも正答率が58％と低い。また、ＭＨでは正答率がＨよりもかえって高くて78％で、ＭＬでは59％である。意味が把握しにくい受け身、使役、補助動詞としての「あげる」「もらう」が一つの問題項目に入っているために、勘で回答した者と途中まで理解した者との差が出ず、受験者の実力が反映されにくかったのであろう。

３－４　読解（35項目）

Ａ　試験の内容

問題Ⅰと問題Ⅱに分かれる。

問題Ⅰは、約1200字からなる問題文を読み、その後で各設問に答える問題である。問題項目は四肢選択が８項目、○×で答える正誤形式が５項目、全部で13項目である。問題文は説明文であり、事実についての叙述に限られている。四肢選択の設問の内容は、問題文の空欄に適切な接続詞を入れるもの、単語の意味を問うもの、語句の具体的な意味を問うもの、題を選ぶものなどである。正誤問題は、問題文の内容にあっているものと違っているものを識別する問題である。

問題Ⅱは、約1600字からなる随筆文である問題文を読み、その後で22項目ある設問に四肢選択で答える形式である。設問の内容は、問題文の空欄に適切な接続詞を入れるもの、筆者の心の動きを問うもの、指示語が指す具体的な内容を問うもの、語句の意味を問うもの、問題文の内容を他の言葉で言い換えるものなどである。

Ｂ　全体像

この試験の結果の概略は以下の通りである。

有 効 数：136
総項目数：　35
平　　均：　19.68
標準偏差：　　4.23

最 高 点： 29

最 低 点： 5

試験全体の分析結果を見ると、平均が35点満点で19.68点（56.23%）とほぼ6割であり、受験者の90%以上が15点から26点までの得点圏に含まれる。60%以上の得点となる21点以上の受験者は54人で、全体の40%にあたる。能力試験としては、統計的に見て全体的に問題はないと考えられる。

C　項目分析

問題Ｉの項目を４つの水準に分類すると表８のようになる。番号は項目の番号を示し、正誤問題の番号は括弧に入れた。

水準Ａ：１
水準Ｂ：３、４、８、（９）
水準Ｃ：６、(10)、(11)、(12)、(13)
水準Ｄ：２、５、７

（表８　読解問題Ｉ）

表８から分かるように、問題文が易しすぎたために水準Ｃの項目が多い。特に正誤問題である項目９～13の５項目のうち水準Ｃの項目が４つもある。○×で答えさせる正誤問題は作成しやすいが、偶然性が入る確率も高く、問題形式として難点があろう。

問題文が易しかったと思われるにもかかわらず、項目１が水準Ａであったことはどのように解釈すべきであろうか。この項目は本文中の空欄に入る適切な接続詞を選ぶ問題である。選択肢は、４級、３級、２級、３級に属する接続詞である。正答率は24%と低く、Ｈグループでも正答率は36.1%である。適切な接続詞を入れる問題は、その前後の文を正しく理解し、その上で、順接になるか逆接になるかなどを判断していかなければならない。最初の問題にこのような高度な能力を要求する問題を出すのは試験作成の方法としてあまりよくない。問題文全体が易しかっただけに試験作成上の配慮が必要な部分である。

能力試験問題の難易度としては全体的に問題Ｉの問題文よりもう少し難しいレベルが必要であろう。

問題IIは、水準Ｃの項目が一つもない。つまり、正答率が80%以上の問題がない。全体的な難易度は適当であったと思われるが、水準Ｂの問題が多かったこと

が検討事項である。単なる日本語の能力の問題でなく、その他の要素が入ってきた可能性もある。

水準Ａ：24、27
水準Ｂ：17、19、21、22、23、25、28、29、31
水準Ｄ：14、15、16、20、26、30、32、33、34、35

（表９　読解問題Ⅱ）

正答率が25％以下の水準Ａに分類された項目24は、問題文にある「そんなこと」が具体的には何を指すかを問う問題である。問題が難しいために統計的にはいい結果が出なかったが、選択肢を易しくすることでよい問題になる可能性はある。

項目27は、問題文中にある語句の意味を問う問題である。誤答である選択肢ａを回答したものが圧倒的に多く（67.6％）、特にＨグループでは75.0％の者が選んでいる。４グループの中で正答を選んだ者の割合が一番多いのは最下位のＬグループで、その結果、弁別指数がマイナスになっている。これは全体の成績上位者がこの問題に限っては下位者より出来が悪かったということである。考えられる要因としては、未習語があるために、成績上位の者ほど既習語の類推からかえって誤答を選んでしまったということがある。

５　問題点

今回の基礎研究では、信頼性・妥当性の高い日本語能力試験を作成するために、まず昨年に実施されたマレーシア政府派遣学部留学生現地予備教育修了試験の試験結果を統計的に集計した。正答率と弁別指数から問題項目を水準別に分類し、「文字・語彙」と「文法」の試験に関しては、さらに『日本語能力試験出題基準』との照合を行った。これまでのマレーシア政府派遣学部留学生現地予備教育修了試験の問題作成では、過去の問題形式や問題数のみが基準として提示され、問題内容の基準は問題作成者に任せられていたが、今回の分析の結果、ほぼ「日本語能力試験２級」レベルの内容になっていたことが明らかになった。

これまでの試験作成方法を問題項目の分析に基づいて検討した結果、いくつかの問題点を確認することができた。その問題点を以下に挙げてみる。

問題形式に関しては、正誤問題は四肢選択に比べて偶然性の関与する確率が高く、信頼性の点で問題があることがこの試験からも確認された。能力試験の形式としては、四肢選択に統一すべきであろう。しかし、四肢選択問題であるにもか

かわらず、実質的には二肢、三肢選択になっていた問題もあった。選択肢の「まよわし」の有効性の問題に十分考慮し、学習者がよく間違えるような点を日頃から記録し、それを選択肢に含めて問題作成をしていく必要がある。

内容に関しては、『出題基準』の１級に分類されている漢字・語彙・文法事項が２級に分類されている漢字・語彙・文法事項よりも必ずしも難しいという結果にはならなかった。要素の試験であればあるほど、試験項目が既習事項であるかどうかが結果に直接反映されるからであろうと考えられる。他機関で日本語教育を受けた学習者に同じ試験を受けてもらい、その結果を比較・検討する必要があろう。

読解、聴解については、そこで使われる語・漢字・文法事項の難易だけでは問題全体の難易度は判断できない。読解の１番目の項目はその例である。読解・聴解で測定するものの難易については、取り上げる素材、題目、文章形式、談話構成などの他に、一般知識を始め、読解・聴解で求められる推測、予測力などを有効に働かせるメカニズムの解明とそれを測定する方策の研究を重ねて行く必要がある。

注

(1) 教育の詳細は小川誠（1995）を参照。

(2) 国際交流基金・日本国際教育協会（1994)。

(3) 外国人日本語能力試験実施委員会企画小委員会（1993)。

参考文献

池田央『テストの科学』日本文化科学社、1992年。

石田敏子『入門日本語テスト法』大修館書店、1992年。

小川誠「マラヤ大学予備教育課程における日本語教育」『日本語教育』85号、1995年、151-159頁。

外国人日本語能力試験実施委員会企画小委員会『日本語能力試験の概要1992年版』国際交流基金、1993年3月。

国際交流基金・日本国際教育協会『日本語能力試験　出題基準』国際交流基金、1994年11月。

竹谷誠『新・テスト理論』早稲田大学出版部、1991年。

横田淳子・伊東祐郎・西郡仁朗「学部留学生の日本語能力試験開発のための基

礎研究(1)」『東京外国語大学留学生日本語教育センター論集』第21号、1995年、29-48頁。

本研究は平成6年度東京外国語大学教育研究学内特別経費による助成を受け行われた。なお、試験結果のデータ処理に関しては今村大介氏（東京学芸大学大学院生）の協力を得た。

A Preliminary Study for Development of the Japanese Proficiency Test for Pre-College Students (2)

YOKOTA Atsuko & ITO Sukero

The purpose of this research series is to obtain useful information for developing the Japanese proficiency test for pre-college students. Last year, the authors analyzed the result of the term-end Japanese tests conducted in December 1993, at the Japanese Language Center for International Students, Tokyo University of Foreign Studies. This year, they examined the result of the final Japanese test consisting of four parts, i.e, listening, kanji, grammar and reading conducted at the Preparatory Course in Malaya in January 1994.

In addition to the statistical analysis of the result of the tests in all four parts, the items of the kanji and grammar tests are examined by checking each item against the Standards of Levels for the Japanese Language Proficiency Test.

It is found out that the level of the Malaya final Japanese test is generally at the second level of the Japanese Language Proficiency Test, even though the results of some discrete test items such as kanji and grammar have shown the influence of whether the test items are already learned by the students or not. It is, therefore, useful to compare the results of the tests with one taken by students of other institution for further investigation.