東京外国語大学
留学生日本語教育センター論集　23：43～65，1997

学部留学生の日本語能力試験開発のための基礎研究（3）

横田　淳子・伊東　祐郎

（1996.10.31　受）

1．研究の意義と目的

学部進学留学生用の日本語能力試験を開発するにあたって、測定すべき日本語能力を明らかにする一つの方法として、1993年度には、文部省学部留学生の予備教育機関である東京外国語大学留学生日本語教育センターの2学期期末試験（以下、「期末試験」）を[1]、1994年度には、マレーシア政府派遣学部留学生現地予備教育の最後に行われる修了試験（以下、「修了試験」）を[2]研究の対象とし、それぞれの試験について試験結果を分析した。試験の項目分析のほかに、「期末試験」の受験者には「日本語能力試験1989年2級」を受験してもらい、文法、読解、聴解の試験に関して、「期末試験」の試験結果と「日本語能力試験1989年2級」の試験結果との相関関係を調べた。「修了試験」の試験問題の漢字、語彙、文法事項に関しては『日本語能力試験出題基準（以下、『出題基準』）』[3]と照合し、試験内容のレベルを分析した。その結果、両方の試験とも、総体的に「日本語能力試験2級」にほぼ相当することが確認された。しかし、1級に分類されている漢字、語彙、文法事項の試験項目が2級に分類されている漢字、語彙、文法事項の試験項目より必ずしも正答率が低いという結果にはならず、これは、試験項目が受験者の教育機関で教えられているかどうかが試験結果に大きく反映しているためであると考えられた。

学部進学留学生用の日本語能力試験は、受験者が現時点で大学学部において勉学できる日本語力を有しているかどうかを測定すると同時に、勉学に必要な日本語力を今後も習得していく可能性を有しているかをも測るものでなければならない。今まで学んだことがどれだけ習得されているかだけを測定するのでは不十分なのである。その意味では、特定の教育機関の教育内容が反映される試験であってはならない。学習者が国外・国内どこの予備教育を受けようとも、その予備教育の内容に関係なく、学習者の潜在能力をも含めた日本語能力を正しく測定できるものであることが望ましい。試験が特定の教育内容に片寄らず、一般性のあるものであるか否かを検証するためには、異なる教育機関で日本語教育を受けた学

習者に同一の試験を受験させ、その結果を比較・検討し、試験内容を吟味することが有効であると考えられる。

以上の点から、1995年度は、国内・国外二つの予備教育機関の在学生に同一の試験を受験してもらい、その結果を比較・検討することとした。また、試験問題の漢字、語彙、文法事項を、従来と同様に『出題基準』と照合し、試験内容のレベルを分析するとともに、両教育機関の教科書を参考に、既習事項であるかどうかを調べることとした。

2．研究方法

2-1　分析資料

分析の対象としたのは、新たに国外の予備教育機関の修了試験として作成された日本語の試験である。聴解、文字・語彙、文法・読解の3種類に分かれているが、試験結果の分析にあたっては、以下の6種に分類し、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。

漢字の2つの試験以外は記号または○×で解答するものである。

1	文法	20問
2	読解	21問
3	聴解	25問
4	漢字の読みを書かせる問題	40問
5	漢字を書かせる問題	40問
6	語彙	20問

試験の実施時期と受験者数は以下の通りである。

	国外	国内
実施時期：	1995年1月	1996年1月
受験者数：	125名	41名

2-2　分析プログラム

従来と同様に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 国外の学生のID番号950001～950125、国内の学生のID番号951001～951041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、国外の受験者の統計量（「国外」）、国内の受験者の統計量（「国内」）および国外・国内あわせた受験者の統計量（「MIX」）を求め、それぞれ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）正答率と弁別指数
- 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）

次にSPSSを使って、各試験の信頼性係数を求めた。

2-3 分析方法

個々の試験項目を統計的数量の面と内容面から分析した。

統計的分析のためには、二つの集団（国外・国内）別に各問題項目に関して正答率と弁別指数を出し、『日本語能力試験の概要（試験結果の分析）』で用いられている基準に従って以下の4つの水準に分類した[4]。

【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。

内容面の分析のためには『出題基準』を参照し、それぞれの試験項目の文法、漢字、語彙がどのレベルに属する事項であるかを検討した。文法については、『出題基準』の1級および2級の［文法的な〈機能語〉の類］、3、4級の［文法事項］と［表現意図等］、［1、2級語彙表］［3、4級語彙表］を使い、漢字と語彙に関しては、［1級漢字表］［2級漢字表］および［1、2級語彙表］を使った。

また、国外・国内両機関の教科書を参考に、試験項目がそれぞれの受験集団にとって既習事項であるかどうかを調査した。

3．試験結果の分析

3-1　全体像

【文法】

	MIX	国　外	国　内
有効数	166	125	41
総項目数	20	20	20
平均	9.46	9.30	9.93
標準偏差	2.59	2.51	3.21
最高点	16.00	14.00	16.00
最低点	3.00	3.00	4.00
信頼係数	.44	.44	.58

【読解】

	MIX	国　外	国　内
有効数	166	125	41
総項目数	21	21	21
平均	11.54	10.80	13.78
標準偏差	3.61	3.20	4.41
最高点	20.00	18.00	20.00
最低点	4.00	4.00	5.00
信頼係数	.65	.54	.74

【聴解】

	MIX	国　外	国　内
有効数	167	125	42
総項目数	25	25	25
平均	14.50	13.48	17.55
標準偏差	3.63	3.26	3.93
最高点	22.00	21.00	22.00
最低点	4.00	4.00	11.00
信頼係数	.58	.48	.46

【漢字／読み】

	MIX	国　外	国　内
有効数	167	125	42
総項目数	40	40	40
平均	28.09	30.83	19.93
標準偏差	7.8	5.58	8.64
最高点	39.00	39.00	37.00
最低点	3.00	14.00	3.00
信頼係数	.91	.81	.92

【漢字／書き】

	MIX	国　外	国　内
有効数	167	125	42
総項目数	40	40	40
平均	24.44	27.10	16.55
標準偏差	8.5	5.98	10.44
最高点	37.00	37.00	36.00
最低点	0.00	14.00	0.00
信頼係数	.92	.83	.95

【語彙】

	MIX	国　外	国　内
有効数	167	125	42
総項目数	20	20	20
平均	11.75	12.02	10.95
標準偏差	2.89	2.48	4.16
最高点	17.00	17.00	16.00
最低点	3.00	7.00	3.00
信頼係数	.53	.37	.73

3-2　文法

問題Ⅰ（項目１～10）と問題Ⅱ（項目11～20）の２種類から構成され、答えはすべて四肢選択による。問題Ⅰは、最大50字ほどの文の中の空欄に適切な語句を選んで入れる問題で，選択肢の語句は、接続詞、助詞相当句、文末表現などである。問題Ⅱは、一つまたは二つの文が与えられ、その文の中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。

次の表は、各問題項目の選択肢（ａ～ｄ）および問題文が『出題基準』のどの級に属しているか（？印はどの級か判断しかねたものを示す）を表している。また、それぞれの選択肢を選んだ受験者数を国内・国外別にパーセントで、さらに、各教育機関の教科書に提出されているか否かを○と×で記している。問題Ⅰ（項目１～10）の選択肢はすべて問題通り表示してあるが、問題Ⅱ（項目11～20）では、問題文、選択肢とも長いので、一部を例示するにとどめてある。選択肢ａ～ｄの前の＊印は正答を表している。各問題の項目分析の結果は国外・国内別に（　）内にＡ～Ｄの水準で示している。

問題Ⅰでは、全選択肢40語句中２級に属するものが31語句、３級に属するものが４語句、１級に属するものが３語句である。また、正答の選択肢だけを見ると、10語句中９語句が２級の語句である。残りの１語句「ようなら」は、「よう」「なら」それぞれは３級に属する語句であるが、「よう」「なら」を一緒にした用法は３級より上のレベルになると考えられる。

問題Ⅱでは、問題文の下線部分が文全体に及んでいる項目もあり、単純に『出題基準』と照合することができない。文法事項だけでなく語彙も『出題基準』と照合し、総合的に級を判断した結果、10文すべて２級に属すると考えられた。選択肢では問題文の下線部分を言い換えたものであるため、２級、３級だけでなく４級に属するものもある。また、項目20のように言い換えが文全体に及ぶため、級が判断できなかったものもある。

表を検討すると、正答が未習の語句である場合は、当然のことながら正答率が低くなる（項目１、３、６、７、８）。項目３、６、７の正答は国外・国内両集団にとって未習の事項であったが、項目１と項目８の正答の語句は、片方の集団は未習であるがもう一方の集団は既習ということで、正答率が大きく異なっている。文法問題では問題項目が既習か否かということの正答率に与える影響がいかに大きいかがわかる。

問題Ⅱの場合は、半分の５項目（項目12、13、17、19、20）が国外・国内とも

に統計的観点からはよい項目と分類される。これらの項目では国外・国内の集団の選択肢の選び方にも似た傾向が見られる。特に項目17、19では各選択肢を選んだ受験者のパーセントが大変近い数字になっている。同じ文法の問題でも、問題Ｉより問題IIの形式の方が未習か否かの影響が少なく、能力試験の形式としては適切であると言えるだろう。

【項目１】	＊a のあまり	b のついでに	c をこめて	d をめぐって
出題基準	２級	２級	２級	２級
国外（D） 教科書	88.8% ○	8.0% ×	2.4% ○	0.8% ×
国内（A） 教科書	4.9% ×	29.3% ×	34.1% ○	26.8% ○

【項目２】	a ところで	b ところに	＊c 以上は	d 上に
出題基準	３級	２級	２級	２級
国外（D） 教科書	24.0% ○	9.6% ○	55.2% ○	11.2% ○
国内（D） 教科書	26.8% ○	12.2% ○	31.7% ○	29.3% ○

【項目３】	a 思うからといって	b 思うまでもなく	c 思うどころか	＊d 思いつつ
出題基準	２級	１級	２級	２級
国外（B） 教科書	28.8% ○	7.2% ×	37.6% ○	26.4% ×
国内（A） 教科書	12.2% ×	7.3% ×	68.3% ○	12.2% ×

【項目4】	a ばかりか	b だけか	＊c どころか	d のみで
出題基準	2級	2級	2級	2級
国外（D） 教科書	24.8% ○	0.8% ○	73.6% ○	0.8% ○
国内（D） 教科書	41.5% ○	9.8% ○	46.3% ○	0.0% ○

【項目5】	a あげく	＊b 上で	c 末に	d が最後
出題基準	2級	2級	2級	1級
国外（C） 教科書	9.6% ○	88.8% ○	0.8% ○	0.8% ×
国内（D） 教科書	0.0% ×	73.2% ○	26.8% ○	0.0% ○

【項目6】	a ばかりで	b そばから	＊c と思ったら	d とみると
出題基準	?	1級	2級	?
国外（A） 教科書	79.2% ○	4.0% ×	0.8% ×	16.0% ×
国内（A） 教科書	82.9% ○	2.4% ×	4.9% ×	9.8% ×

【項目7】	a かぎりは	＊b ようなら	c ものなら	d かと思うと
出題基準	2級	3級	2級	2級
国外（A） 教科書	48.8% ○	12.0% ×	34.4% ×	4.8% ○
国内（A） 教科書	29.3% ○	19.5% ×	51.2% ×	0.0% ×

【項目８】	＊a　ほかない	b　ものだ	c　ことにほかならない	d　にきまっている
出題基準	２級	２級	２級	２級
国外（A） 教科書	5.6% ×	14.4% ○	44.0% ○	36.0% ○
国内（C） 教科書	80.5% ○	2.4% ○	17.1% ○	0.0% ○

【項目９】	a　いったい	b　きっと	c　やはり	＊d　まさか
出題基準	２級	２級	２級	２級
国外（A） 教科書	24.8% ○	33.6% ○	20.8% ○	20.8% ×
国内（D） 教科書	4.9% ×	29.3% ○	7.3% ○	58.5% ×

【項目10】	a　そうですね	b　らしいですね	c　ものですね	＊d　ことですね
出題基準	３級	３級	２級	２級
国外（B） 教科書	3.2% ○	15.2% ○	32.0% ○	49.6% ○
国内（D） 教科書	4.9% ○	9.8% ○	19.5% ○	65.9% ○

【項目11】	下線箇所	a　よくある	b　あるはずだ	c　かもしれない	＊d　ないだろう
出題基準	２級	４級	３級	３級	３級
国外（C） 教　科　書	○	0.0% ○	0.8% ○	1.6% ○	97.6% ○
国内（C） 教　科　書	○	0.0% ○	0.0% ○	4.9% ○	95.1% ○

【項目12】	下線箇所	＊a てみたら	b たのに	c とすぐ	d 場所から
出題基準	2級	3級	3級	?	2級
国外（D）教科書	○	32.0% ○	10.4% ○	50.4% ○	7.2% ○
国内（D）教科書	○	29.3% ○	14.6% ○	53.7% ○	2.4% ○

【項目13】	下線箇所	a にたりない	＊b でしかない	c とは限らない	d だけのことはある
出題基準	2級	?	2級	?	2級
国外（D）教科書	○	14.4% ○	51.2% ○	12.0% ○	22.4% ○
国内（D）教科書	○	4.9% ○	61.0% ○	14.6% ○	19.5% ○

【項目14】	下線箇所	a たくなくなる	b ～ようとしても～ない	＊c ～ないでいる～できない	d ～時間はない
出題基準	2級	3級	3級	3級	4級
国外（D）教科書	○	4.8% ○	12.0% ○	82.4% ○	0.8% ○
国内（C）教科書	○	0.0% ○	0.0% ○	100.0% ○	0.0% ○

【項目15】	下線箇所	＊a やめましょう	b やめることにしますか	c やめられません	d やめるつもりですか
出題基準	2級	4級	3級	3級	3級
国外（A）教科書	○	17.6% ○	12.0% ○	7.2% ○	63.2% ○
国内（D）教科書	○	43.9% ○	19.5% ○	7.3% ○	29.3% ○

【項目16】	下線箇所	a ことはない	b ないことになる	c のをやめる	*d なければならない
出題基準	2級	2級	3級	3級	3級
国外（D）教科書	○	12.0% ○	9.6% ○	2.4% ○	76.0% ○
国内（C）教科書	○	2.4% ○	4.9% ○	2.4% ○	87.8% ○

【項目17】	下線箇所	*a 場合でも	b かどうか	c ～ても～ても	d 大いにある
出題基準	2級	2級	3級	3級	2級
国外（D）教科書	×	48.0% ○	16.8% ○	14.4% ○	20.8% ○
国内（D）教科書	○	43.9% ○	17.1% ○	12.2% ○	26.8% ○

【項目18】	下線箇所	a ～しても、～	b ～すればするほど、～	c ～したら、～	*d ～するまで、～
出題基準	2級	3級	2級	3級	3級
国外（A）教科書	○	36.8% ○	47.2% ○	4.0% ○	12.0% ○
国内（D）教科書	○	26.8% ○	36.6% ○	2.4% ○	31.7% ○

【項目19】	下線箇所	a ことがある	b ことはできない	*c ～られない	d ～てしまう
出題基準	2級	3級	3級	3級	3級
国外（D）教科書	×	12.8% ○	26.4% ○	54.4% ○	6.4% ○
国内（D）教科書	○	14.6% ○	26.8% ○	56.1% ○	2.4% ○

【項目20】	下線箇所	a ～し	*b 部分もあるが	c ～が	d いないと思う
出題基準	2級	?	?	?	?
国外（D） 教科書	○	5.6% ○	37.6% ○	36.0% ○	20.0% ○
国内（D） 教科書	○	2.4% ○	46.3% ○	46.3% ○	2.4% ○

3-3 読解

問題Ⅰ（項目1～12）は、人生における道草に関する、800字ほどのまとまった文を読んで、文中の語句の表している意味、文全体の意味、空欄に入れることばなどを四肢選択で、あるいは語群の中から適切な語を選んで答えさせるものである。

問題Ⅱ（項目13～21）は、人間の脳の情報蓄積能力についての1400字ほどの長文を読んで、文中の空欄に接続詞を入れさせたり、内容に関する質問に答えさせたりするもので、すべて四肢選択である。

各問題項目の正答率（%）と弁別指数は以下の通りである。

項目	国外			国内		
	正答率%・弁別指数・水準			正答率%・弁別指数・水準		
1	69.6	0.319	D	75.6	0.583	D
2	62.4	0.644	D	75.6	0.409	D
3	41.6	0.220	B	73.2	0.061	B
4	44.8	0.372	D	85.4	0.250	C
5	39.2	0.371	D	48.8	0.030	B
6	34.4	0.520	D	68.3	0.667	D
7	75.2	-0.037	B	80.5	0.583	D
8	44.0	0.581	D	63.4	0.826	D
9	44.0	0.282	B	39.0	0.303	D
10	19.2	0.336	A	48.8	0.561	D

11	33.6	0.310	D	34.1	0.735	D
12	55.2	0.644	D	70.7	0.583	D
13	48.8	0.400	D	70.7	0.235	B
14	85.6	0.235	C	87.8	0.250	C
15	38.4	0.430	D	58.5	0.386	D
16	75.2	0.261	B	80.5	0.500	D
17	61.6	0.198	B	63.4	0.152	B
18	56.8	0.434	D	68.3	0.409	D
19	86.4	0.175	C	82.9	0.417	D
20	45.6	0.611	D	65.9	0.659	D
21	18.4	0.275	A	36.6	0.386	D

問題はどれも文の内容に関するもので、単純に単語の意味や文法的知識を問うものはなかった。このため、ある語または漢字・文法事項が既習か否かといったことが試験結果にほとんど関係なかったと言える。国外・国内どちらも水準Cとなった項目14は、逆接関係を読み取って、逆接の接続詞「しかし」を入れる問題である。誤答の選択肢である「ところが、そして、すなわち」を選んだ受験者はわずかであった。国外・国内どちらも水準Bとなった項目3は、考え過ぎるとかえって迷う問題である。選択肢を工夫することによって水準Dになると思われる。同じく国外・国内どちらも水準Bとなった項目17は「前者・後者」の指すものを問う問題である。正答率が比較的高くなり、低得点者も高得点者に近い割合で正答を選んでいるために弁別指数が低くなった。項目4は国内で水準Cとなり、易しすぎる問題となったが、問題文の内容を問う問題であり、特定の語彙・漢字・文法事項が既習であることの結果、二つの集団で差が生じたとは考えられない。国外で水準Aとなった項目21は、文章の最後に入る文を選ぶ問題である。問題文全体の主旨が読み取れれば答えられるはずであるが、国外の場合は、低得点者の多く（70%以上）と高得点者の半数近くが常識的には妥当な答えとなる誤答の選択肢を選んでしまった結果、正答率・弁別率ともに低くなったと考えられる。

全体的に水準Dが多いことからも、問題文のレベルは妥当だったことがわかる。ただし、項目7～12の6問は、四肢選択ではなく共通の11の語群から語を選ぶもので、実際は二肢選択になっていたものが多かった。項目7は「作家」と「おとぎ話」の二肢選択の問題となったが、結果的に国外では低得点者の方が高得点者

よりも正答率が高くなり、弁別率がマイナスになった。問題形式としては、1項目ずつ四肢選択の形にしたほうがやはりよかったのではないだろうか。

3-4 聴解

出題形式の異なる3つの問題群（Ⅰ～Ⅲ）から構成されている。解答はすべて選択式である。問題Ⅰ（項目1～9・四肢選択）は、イラストやグラフなどの視覚情報から正しいものを選ぶ問題である。問題Ⅱ（項目10～15・四肢選択）は、音声情報のみを聞いて答える問題で、選択肢も音声で与えられる。問題Ⅲ（項目16～25・正誤）は2問あり、それぞれ比較的長い音声情報を聞いた後で、5つの音声内容の適否を○×で答えるものである。

各問題項目の正答率（%）・弁別指数は以下の通りである。

項目	国外			国内		
	正答率%	弁別指数	水準	正答率%	弁別指数	水準
1	84.0	0.174	C	95.2	0.091	C
2	11.2	0.033	A	11.9	0.091	A
3	90.4	0.086	C	95.2	0.182	C
4	33.6	-0.079	B	45.2	0.455	D
5	43.2	0.461	D	64.3	0.364	D
6	52.0	0.433	D	71.4	0.364	D
7	41.6	0.372	D	61.9	0.000	B
8	62.4	0.435	D	59.5	-0.091	B
9	57.6	0.197	B	78.6	0.455	D
10	44.8	0.433	D	76.2	0.182	B
11	79.2	0.263	B	85.7	0.364	D
12	36.0	0.220	B	73.8	0.182	B
13	57.6	0.406	D	85.7	0.182	C
14	40.0	0.520	D	69.0	0.455	D
15	63.2	0.496	D	57.1	0.636	D
16	64.0	0.409	D	88.1	0.091	C
17	61.6	0.258	B	83.3	0.364	D

18	19.2	-0.026	A	33.3	0.818	D
19	63.2	0.317	D	88.1	0.182	C
20	78.4	0.260	B	88.1	0.273	C
21	65.6	0.378	D	76.2	0.455	D
22	52.8	0.316	D	76.2	0.091	B
23	24.8	0.217	A	57.1	0.364	D
24	68.0	0.288	B	71.4	0.000	B
25	53.6	0.463	D	61.9	0.182	B

聴解問題は質問内容から判断して言語的要素というより聴解活動に関するものから構成されていると言ってよいだろう。従って、各問題項目の水準から問題項目及び選択肢の内容について検討してみる。

まず、項目２が国外・国内ともに水準Ａで、双方にとって難しい問題であったことが分かる。「左の棚」「ドアの後ろ」「上から３段目」「２段目」などから正しい位置を認識する問題であるが、国内外を問わず位置を問う問題は難しいのであろうか。また、イラストが分かりにくかったことも影響しているのかもしれない。項目４では国外の正答率は33.6％と低い。船での移動の場合、風向によって速度が変わることを予備知識としてもっていなければならない。イラストが島の地図となっているが、東西南北の把握が難しかったのであろうか。項目12は国内外ともに同水準であるが、国外の正答率が国内の半分である。海外旅行に伴う時差の経験の有無が結果に現れた可能性も考えられる。

項目18の正答率は国内外ともに低い。「サマータイム導入で、スポーツやレジャーをもっと楽しむ必要がでてきます」の内容の適否を判断する問題である。正答は×であるが、○を選んだものが大多数である。聴解テクストでは「サマータイム導入で、スポーツやレジャーをもっと楽しむことができるが、国民の間では、かえって仕事が増えそうという反対論がある」と述べられている。サマータイム導入でできることに焦点を当てるならば、○であろうし、反対意見に焦点を当てるならば×であろう。設問内容があいまいで不適切であったことが考えられる。項目23では国外の正答率が国内の半分以下と低い。「現在、この国ではゴミを５種類に分けて捨てることになっています」の適否を判断する問題である。正答は×である。テクストでは「来年度から５種類に分けて捨てることになっています」で、まだ実際には実行に移されていないことが分かる。したがって、設問

内容は現在の状況の説明になっているので明らかに×となる。詳細に聞き取ることが求められる問題である。

項目８と15を除く他のすべての項目で、国内受験者の正答率が国外のそれより高い。このことから、聴解力においては国内という学習環境面など有利な点が影響していることが考えられよう。しかし、問題ＩとIIは四肢選択からなるが、国内では15項中７項目（項目１、３、４、８、９、10、13）において、選択率０％の選択肢があった。特に項目13では２つもあり、二肢選択問題と変わらない。選択肢の作り方を再考し、国外・国内という要因によってテクスト内容の難易度が左右されないよう注意しなければならない。

テクストタイプでは男女の会話形式が15項目中11項目と比較的多い。それ以外は、説明・講義等からなる。聴解問題では、トピックに関する言語的知識をはじめ、社会文化的知識や内容・背景知識等が必要となる。項目数が限られているので大学生活に関連した、或いは大学入学後に必要とされる聴解テクスト及び課題をより多く出題することで内容的妥当性を高めていく必要があると思われる。

3-5　漢字／読み

漢字の読み方をひらがなで書かせる問題で40項目からなる。漢字１字を送り仮名つきで１項目として取り扱うものと、２字あるいは３字からなる熟語を１項目として扱うものとがある。すべての漢字は、短文の中で出題されている。

全体像から漢字及び語彙に関する平均点は、国内が国外より低い傾向にあることが分かる。特に「漢字／読み」と「漢字／書き」の平均点の差が大きい。

次の表は、問題項目を水準別に分類したものである。この表に基づいて、問題文が国外・国内の教育機関で使用されている教科書に提出されているか否かを調べ、既習あるいは未習漢字・語彙と成績の関係を考察してみる。なお、項目の後の（　）内は、出題語彙と『出題基準』のレベルを表す。

	国　　外	国　　内
水準Ａ	14,16,27	2,3,6,14,16,18,21,26,27,30,33,35,37
水準Ｂ	8	

水準C	1,4,5,7,9,10,12,17,19,22,23,24,25,28,29,31,32,35,36,37,38	10,17
水準D	2,3,6,11,13,15,18,20,21,26,30,33,34,39,40	1,4,5,7,8,9,11,12,13,15,19,20,22,23,24,25,28,29,31,32,34,36,38,39,40

まず、水準Aに分類された正答率25%以下の項目について、教科書の語彙リストと照合した結果、国外では項目14「省みる（1級）」は提出されている。項目16「図る（1級）」、項目27「険しい（2級）」は、語彙としては提出されていないが、前者の場合は「図（書館）」「(地）図」で、後者の場合は「(危）険」「(保）険」で提出されている。

国内では項目18「抱く（2級）」のみが教科書に提出されている。その他の問題項目は教科書では未提出語彙である。しかし、未提出語彙であっても漢字そのものが、別の形で提出されているもの（項目2，6，26，27，30，35）がいくつかある。その例として「犯す（1級）」は「犯（人）」、「定める（1級）」は「定（規）」、「費やす（1級）」は「(旅）費」、「疲労（1級）」は「疲（れる）」と「(苦）労」等で提出されている。国外と同水準の項目14、16は国内では未提出、項目27は「(保）険（証）」で提出されている。

水準B、C及びDに分類された語彙は、国外・国内ともに語彙リスト及び例文等で基本的には提出されている語彙ばかりである。

漢字読みに関しては、国内における未習問題項目数が成績に影響を与え、問題の難易度にも多少関与しているのではないかと思われる。同一漢字であっても音読みと訓読みの双方が学習されているか否か、また、どのような脈絡で、どのような熟語、関連語として導入されているかによっても解答に影響がでるものと考えられる。

3-6　漢字／書き

文中のひらがな部分を漢字に書き換えさせる問題で40項目からなる。漢字の再生能力に関する問題であるため、熟語であっても構成漢字1字が採点対象単位になっている。

	国　　外	国　　内
水準A	20,21,25,28	1,2,20,21,28,29,32,35,36,39
水準B		
水準C	5,6,7,9,11,13,14,17,22,23,24,30,31	
水準D	1,2,3,4,8,10,12,15,16,18,19,26,27,29,32,33,34,35,36,37,38,39,40	3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,22,23,24,25,26,27,30,31,33,34,37,38,40

読み問題同様、水準Aに分類された正答率25%以下の項目について、教科書の語彙リストと照合した結果、まず国外では項目25「実る（2級）」のみが未提出語彙であった。しかし、漢字「実」は音読み熟語「実（家）」「実（感）」で提出されている。一方、国内では、項目29「疑もん（2級）」を除くすべてが未提出語彙であった。語彙としてではなく漢字単独としては、項目28「妨げる（2級）」は「妨（害）」で、項目39「報こく（2級）」は「報（道）」で提出されている。水準Bに分類される項目は国外・国内ともになかった。水準C及びDに分類された語彙は、国外・国内ともに語彙リスト及び例文等で基本的には提出されている語彙ばかりである。

提出語彙であるにもかかわらず、水準Aに分類された項目をみてみると、国外では「犠牲（1級）」「妨げる（2級）」が、国内では「疑（問）」が挙げられる。読めても漢字を再生する場合には、正確な画数や筆跡などが求められる。もちろん、これにより実力が評価できるとも言えるが、字形の正しさの評価に主観的要素がかなり入ることになる。再生技能面からみると漢字圏の受験者に有利になってしまうのではないだろうか。理解語彙と使用語彙を区別するように、漢字再生問題試験の場合、理解漢字と再生漢字を区別し、語彙の選択には十分注意する必要がある。併せて、漢字能力とは何かを明確にしておくことも大切になろう。

3-7　語彙

漢字語彙に関する問題で20項目からなる。文中の空欄に適切な語彙を四肢選択で選ぶ形式である。

	国　　外	国　　内
水準A	5,10,14	5,7
水準B	17,20	11,20
水準C	2,6,7,8,9,12,15	8,12
水準D	1,3,4,11,13,16,18,19	1,2,3,4,6,9,10,13,14,15,16,17,18,19

水準Aに分類された正答率25%以下の項目の正答語彙について、教科書の語彙リストと照合してみた。国外では項目5（返却）、項目10（突破）、項目14（審査）がすべて未提出語彙である。一方、国内では項目5、項目7（紛失）は未提出である。

次の表は、水準Aに分類された問題項目の選択肢とその『出題基準』のレベル、選択率、教科書での提出の有無を示している。＊は正答、？は該当級なしを表す。

【項目5】「図書館で借りた本を_____する」

	a 返品	b 返送	＊c 返却	d 返還
出題基準	?	?	?	1級
国外教科書	1.6% ×	80.0% ×	11.2% ×	7.2% ×
国内教科書	2.4% ×	45.2% ×	16.7% ×	33.3% ×

【項目7】「パスポートを_____する」

	a 失望	＊b 紛失	c 遺失	d 失意
出題基準	2級	1級	?	?
国外教科書	0.8% ×	92.8% ○	4.0% ×	2.4% ×
国内教科書	9.5% ×	23.8% ×	45.2% ×	16.7% ×

【項目10】「留学生の数が10万人を_____する」

	a 突入	b 突出	c 突然	＊d 突破
出題基準	?	?	2級	1級
国外教科書	65.6% ×	24.8% ×	3.2% ×	6.4% ×
国内教科書	31.0% ×	23.8% ×	2.4% ○	42.9% ○

【項目14】「論文は教授によって_____される」

	a 診察	＊b 審査	c 検査	d 観察
出題基準	2級	1級	2級	2級
国外教科書	13.6% ×	23.2% ×	28.0% ○	35.2% ○
国内教科書	14.3% ×	31.0% ×	42.9% ○	11.9% ○

上の表から、項目5ではすべて未習語彙であったためか、「返送」の「送」（国内外ともに提出漢字）を頼りに解答したのではないかと考えられる。「返却」は『出題基準』ではいずれの級にも属さない語彙である。

項目7では正答語彙が国外で既習であったためか正答率が高い。一方、国内では未提出のためか4分の1の選択率である。語彙としては1級レベルである。

項目10では国外の教科書では未提出であり、正答率は低い。国内では正答が既習語彙であったために正答率が高いと言えよう。

項目14は、国外・国内ともに正答より既習語彙である誤答を選んでいるものが多い。選択肢が類似した語彙で構成される場合は、既習語彙を選ぶ傾向があると思われる。

4．まとめ

現在、多くの学部留学生は日本語学や日本文学などよりも経済・経営や工学系といった分野を大学で専攻している。このような、大学で日本語を伝達の手段と

して使用しようとする学部進学留学生を対象として、入学時に最低限必要な日本語力を有しているかどうか、また、さらに伸びて行く力があるのかどうかを測定・予測できるような能力試験[5]を開発したいと考え、3年にわたって既存の試験の分析を行ってきた。

今回の調査からは、文法・漢字・語彙などの言語要素の試験において、受験者が試験項目を予備教育で学んだかどうかが大きく試験結果に影響していることが明らかになった。能力試験の場合、試験項目が既習であるか否かが直接結果に表れるような試験は好ましくなく、やはり避けるべきであろう。

助詞の問題や漢字の書きなどは、実際に大学で勉学する上でどの程度問題になるのであろうか。大学では、文法事項の知識が直接問われることはなく、その知識を使って、いかに聞き、話し、読み、書きができるかが問われるのである。ワープロでレポートを提出することが一般的になっている現在、漢字の1画がはねているか、突き出ているかなどということを能力試験で測定することに意味があるだろうか。形成的評価を目的としての試験であれば、文法や漢字などの言語要素を直接問い、動詞の活用、助詞などを細かくチェックするのも適当かもしれないが、能力試験という観点からは再考を要するであろう。能力試験としては、できるだけ大学の場で実際に要求されるような機能面を重視した問題構成にして、新しい場面で日本語を操作して伝達を行うことができる能力を測定・評価する試験であることが望ましいと考える。

最後に、学部進学留学生用の日本語能力試験として、以下のような試験の形式や内容を試みとして提言したい。

文法は、空欄に適切な語句を入れる形式よりも、より読解に近いものにはなるが、前後の文脈から判断できるように文の一部または全部を簡単な言葉で言い換える形式のほうがいいであろう。また、文字によるペーパーテストだけでなく、音声を使ったテープによるテストで、文法的なことを測定することも考えられる。

読解は、題材によって受験者に不公平にならないように配慮する必要がある。題材にあまり日本事情的なものや専門分野に関係するものは使えないため、一般的な内容の論説文、随筆になりがちである。しかし、読まなくても常識で判断できるものや受験者がよく知っている内容のものでは読解力を測ることにならない。そのために、著者の気持ちや意見・考えを問う問題になりやすいが、大学での読解を考慮するともっと事実を説明した文を問題文に入れた方がいい。また、題材によるバイヤスを少なくする意味で、長い問題文を2つ、3つ出題するだけでな

く、１段落程度の長さの文章を多岐にわたる分野から多数出題し、その要旨を問うような問題もいいのではないだろうか。

聴解では、日常生活を含めた大学生活全般で学部進学予定者が出会うテクスト場面を設定し、それに関する聴解タスクを出題することが大切となろう。そのためには、一般の大学生のコミュニケーション行動及び聴解行動をまとめ、出題領域を整備することも必要になってくるものと思われる。テクストタイプも大学生活での会話・講義・説明などある程度限定したものにすることが望ましいと考える。

漢字・語彙に関しては、特に既習・未習の影響が試験結果に反映されやすいので、語彙選択には特に注意を要する。漢字の読みについては、文脈から推測できるようなもの、また音読み・訓読み能力を測定できるような構成にするとよいだろう。形式としては、読解問題のなかに組み込んでしまうのも一方法である。解答は実際に書かせるものより選択式にしたほうが、記入ミスなどを排除できる。漢字の書きに関しては、能力試験では必要ないと考える。むしろ、漢字認識力と理解力、使用力を語彙問題の形式のような文脈のなかで推測させ、本来の運用能力を測定できる形式にした方がよいのではないだろうか。

註

（1） 横田・伊東・西郡（1995）参照。

（2） 横田・伊東（1996）参照。

（3） 国際交流基金・日本国際教育協会（1994）。

（4） すべての問題が四肢選択の問題でないことから、このような分類のしかたには問題があるとも思われたが、試験内容を検討する手掛かりとするためにすべての問題項目を対象に項目分析を行った。

（5） 予測という面を重視すると、予測試験（prognostic test）、適性試験（aptitude test）という名称も考えられるが、本論では、「ある個人のもつ語学力や学習言語に関する知識をその個人が受講したコースや使用した教科書の内容とは関係なく試すテストで、一般に、仕事やコースなど、ある特定の領域で必要とされる語学力と結び付けて実施される」（石田1992、p.18）試験という意味で、「能力試験」という名称を使用している。

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。

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（試験結果の分析）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。

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。

横田淳子・伊東祐郎「学部留学生の日本語能力試験開発のための基礎研究（２）」『東京外国語大学留学生日本語教育センター論集』第22号、1996年。

横田淳子・伊東祐郎「学部進学予定者のための日本語能力試験開発――文法試験を中心に――」『第５回小出記念日本語教育研究会論文集』1996年。

本研究は平成７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。昨年度に続き、試験結果のデータ処理では今村大介氏（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 (3)

YOKOTA Atsuko & ITO Sukero

In the last two years, the authors have analyzed the results of the term-end Japanese language tests conducted in the preparatory education for college at the following two institutions : the Japanese Language Center for International Students, Tokyo University of Foreign Studies (December 1993) ; and the Preparatory Course in Malaya (January 1994).

This year, for the purpose of examining how each test item gives an effect on the test results depending upon whether or not such item has already been learned by the students, one of the existing tests conducted at the end of preparatory education was used for this research. The test is made up of the following six sections : 1. Grammar ; 2. Reading Comprehension ; 3. Listening Comprehension ; 4. Kanji Reading ; 5. Kanji Writing ; and 6. Vocabulary. The test was taken by the two groups studying at the two different institutions : 125 students outside of Japan (ABROAD) and 41 students in Japan (JAPAN), and the test results of the two groups were compared and statistically analyzed.

Item analyses (item difficulty and item discrimination) were conducted for all test items. The results showed that the patterns of performance were different between the two groups. In order to investigate the reasons for the difference of performance, each item was checked against the Standards of Levels for the Japanese Language Proficiency Test and the textbooks of the two groups. Consequently, it has been found out that the level of the test is generally at the second level of the Japanese Language Proficiency Test but the item difficulty and discrimination tend to vary depending upon whether or not the examinee has studied the item at the institution beforehand.