

研究ノート

学部留学生向け日本語プレースメントテストの開発について

京 祥太郎*1

キーワード：留学生、Web 日本語プレースメントテスト、細目表、試行テスト、項目応答理論

1 はじめに

至誠館大学東京キャンパス（以下、東京キャンパス）では、これまで学部留学生を対象に「入学前オリエンテーション」として、入学許可者に入学前に1日登校してもらっている。当日実施したアンケート調査や日本語プレースメントテストなどのレディネス調査の結果をもとに、1年次の科目「基礎ゼミⅠ」「日本語Ⅰ」「日本語特講Ⅰ」のクラス分けを行い、今年度はある程度の成果があった。入学前教育をしっかりと取り組んでいなかった、あるいは、日本語プレースメントテストを受けなかった新入生の多くに、入学後に修学困難になっている現状が見受けられた。

そこで、本稿では、日本語プレースメントテストの開発に向け、まず、試行テストのための細目表を作成し、「基礎ゼミⅠ」の学生に試行テストを実施した。試行テストの結果について項目分析を行い、妥当性や信頼性について検討する。さらに、どのような日本語プレースメントテストが学部留学生にとって効果的なのか、今後の展開方法と検討課題について報告する。

2 プレースメントテスト開発の背景

東京キャンパスの留学生には、日本語能力試験（以下、JLPT）のN1もしくはBJT ビジネス日本語能力テスト480点以上を在学中（2年次後期頃まで）に合格できるよう指導をしており、正規の言語科目「日本語特講」を必須科目とし、1年次科目「日本語特講Ⅰ・Ⅱ」でJLPTのN2レベルを、2年次科目「日本語特講Ⅲ・Ⅳ」でJLPTのN1レベルの学習をすることになっている。ここ数年はベトナムやネパールといった非漢字圏

からの留学生が増え漢字を苦手とした学生が多かったことから、1年次科目「日本語特講Ⅰ・Ⅱ」では全クラスで漢字（文字語彙）の学習をすることになっている。その他の日本語の科目として東京キャンパスでは、1年次必須科目「日本語Ⅰ・Ⅱ」、2年次必須科目「日本語Ⅲ・Ⅳ」も開講しており、「日本語Ⅰ～Ⅳ」の到達目標としては「大学で求められるアカデミックジャパニーズを身につける」「JLPTのN2に合格できるようにする」としている。今年度は、概ね1クラス15名程度とし、「日本語特講Ⅰ・Ⅱ」および「日本語Ⅰ・Ⅱ」が10クラス、「日本語特講Ⅲ・Ⅳ」が3クラス、「日本語Ⅲ・Ⅳ」が9クラスを開講している。

東京キャンパスにおける日本語プレースメントテストとしては、Web上で行うTTBJ（筑波日本語テスト集）を各自に受けさせ、その結果によりクラス分けをした（昨年度まではJ-CAT使用）。しかし、TTBJを受ける際のテスト環境を考慮すると、独自の試験を開発する必要があり、今回、東京キャンパスの情報センターから打診され、新入生を対象とした東京キャンパス独自のWeb日本語プレースメントテストを作成することになった。

3 問題項目開発について

3.1 対象者

日本語プレースメントテストを作成するにあたり、まず、試行テストを作成した。試行テストについては、2021年度前期「基礎ゼミⅠ」を履修している51名を対象に「基礎ゼミⅠ」の時間に45分でWebではなく筆記試験として実施した。対象者の国籍は中国37名、

*1 至誠館大学 現在社会学部

ベトナム10名、ネパール2名、モンゴル1名、ウズベキスタン1名である。中国の学生が多いのは、受け持ちの「基礎ゼミⅠ」のクラスが上位から中位クラスのレベルの学生が集まっているクラスであるためである。

3.2 試行テストの内容

表1は、試行テストを作成するための細目表である。伊東（2008）によると、細目表にはテストの目的、構成、出題形式、測定対象となる言語能力などが詳細に記述され、作題者は細目表への記述を通じて、これか

ら測定しようとする能力を具体的にイメージができると述べている¹⁾。試験内容はJLPTのN2レベルの言語知識（文字語彙・文法）とした。言語知識（文字語彙・文法）にした理由は、現在「日本語特講Ⅰ」で主に漢字（文字語彙）の学習をすることになっていること、さらに、「基礎ゼミⅠ」内で実施する時間的都合としてである。また、1年次の到達目標として「JLPTのN2合格」をあげているため、レベルはN2とした。なお、今回は問題項目の妥当性を検討するため、Webではなく筆記試験とし教室で一斉に行うこととした。

表1 細目表

テスト名	試行テスト
測定目標	熟達度を測るテスト
想定される受験者	新入生（基礎ゼミⅠの学生）
試験内容	日本語能力試験（JLPT）N2程度に相当する言語知識（文字・語彙・文法）を有しているかを測る試験
テストの構成	大問1～10：言語知識（文字・語彙・文法）
問題の構成及び出題のねらい	筆記：多肢選択（4択）問題 【文字・語彙】 大問1：漢字読み（漢字で書かれた語の読み方を問う） 大問2：表記（ひらがなで書かれた語が、漢字でどのように書かれるかを問う） 大問3：語形成（派生語や複合語の知識を問う） 大問4：文脈規定（文脈によって意味的に規定される語が何であるかを問う） 大問5：言い換えの類義（出題される語や表現と意味的に近い語や表現を問う） 大問6：用法（出題語が文の中でどのように使われるかを問う） 【文法】 大問7：文法形式の判断（文の内容に合った文法形式かどうかを判断することができるかを問う） 大問8：文の組み立て（統語的に正しく、かつ、意味が通る文を組み立てることができるかを問う）

<p>測定対象 言語知識</p>	<p>【文字・語彙】 大問1：輸入、環境、税金、評価、法律、経営、政治、面接 大問2：調査、混ぜる、怖い、人類、頂上、幸運、出版、甘い 大問3：入居者、日本製、築30年、不燃物、動物園、海水浴、再来年、現住所 大問4：とつくに、出張、(食費が)かかる、早退、まごまご、ける、(落ち)込む、 努力 大問5：当然(当たり前)、びしょびしょ(濡れる)、後日(あとで)、コンクール(大会)、じょじょに(少しずつ)、納得がいかない(満足しない)、消去する(消す)、乾燥する(乾く) 大問6：ユニーク、活気、ぎっしり、手間、確か、(宝くじが)当たる、名所、わざわざ 大問7：一方、(できる)かぎり、につれ、を問わず、といっても、おかげ、しようがない、恐れがある、だけに、たあげく、さえ、といったら、もの、といえば、ほど～はない 大問8：に関する、からでないと、ことだし、た限り、たとたんに、をはじめ、からといって～とは限らない、こととなると、折に、(寂し)げ、(風邪)きみ、になるなんて</p>
<p>解答方法</p>	<p>紙筆</p>
<p>解答時間</p>	<p>45分</p>
<p>項目数と配点</p>	<p>【文字・語彙】 小問数30 計30点 大問1：小問数8 1問1点×5=5点 大問2：小問数8 1問1点×5=5点 大問3：小問数8 1問1点×5=5点 大問4：小問数8 1問1点×5=5点 大問5：小問数8 1問1点×5=5点 大問6：小問数8 1問1点×5=5点 【文法】 小問数17問 計34点 大問7：小問数15 1問2点×12=24点 大問8：小問数12 1問2点×5=10点 <p style="text-align: right;">総合計100点</p> </p>
<p>判定基準</p>	<p>0～49点：JLPT N2 レベル未満と判定 50点以上：JLPT N2 合格レベルと判定 ただし、基準点(文字語彙：9点以上、文法：10点以上)が1つでも下まわった場合はレベルに達していないと判定</p>

3.3 項目分析結果

テストの結果として、文字語彙および文法の代表値を表2に記す。平均点は文字語彙が73点、文法が48点であった。中央値は、文字語彙が73点、文法が44点、最頻値（流行値）は、文字語彙が67点（6人）、文法が44点、50点（6人）であった。

表2 代表値（100点満点）

	文字語彙	文法
平均値	73点	48点
中央値	73点	44点
最頻値	67点（6人）	44点、50点（6人）

さらに、2010年から日本留学試験（EJU）やBJTビジネス日本語能力テストなどで採用されている項目応答理論¹⁾により、今回、テストにおける各設問の正答率と識別力を使って、各設問の良否を検討するための項目分析（GP分析）を行った。文字の項目分析結果を表3に、語彙の項目分析結果を表4に、文法の項目分析結果を表5に記す。横山（2011）によると、識別力の結果が0.4以上なら一般的に識別力が高い問題、0.2以下なら識別力が低い問題と判断されるとされており²⁾、また、正答率は0.2～0.8が望ましい目安とされている。

表3 文字の項目分析結果一覧表

問題番号	受験者数	正答者数	正答率	識別力
1-1	51	30	0.59	0.20
*1-2	51	34	0.67	<u>0.40</u>
1-3	51	48	0.94	0.20
**1-4	51	38	0.75	<u>0.26</u>
1-5	51	50	0.98	0.07
1-6	51	46	<u>0.90</u>	<u>0.27</u>
1-7	51	448	0.94	0.06
1-8	51	49	0.96	0.13

2-1	51	50	0.98	0.07
2-2	51	50	0.98	0.00
2-3	51	466	0.90	<u>0.33</u>
2-4	51	44	0.86	0.20
*2-5	51	37	0.73	<u>0.46</u>
2-6	51	44	0.86	0.27
2-7	51	50	0.98	0.07
2-8	51	49	0.96	0.13

表4 語彙の項目分析結果一覧表

問題番号	受験者数	正答者数	正答率	識別力
3-1	51	47	0.92	0.20
3-2	51	47	0.92	0.20
*3-3	51	30	0.59	<u>0.46</u>
3-4	51	27	0.53	0.00
3-5	51	46	0.90	<u>0.27</u>
3-6	51	46	0.90	0.20
**3-7	51	38	0.75	<u>0.33</u>
3-8	51	44	0.86	<u>0.40</u>
4-1	51	14	0.27	0.13
4-2	51	47	0.92	<u>0.27</u>
4-3	51	42	0.82	<u>0.26</u>
4-4	51	43	0.84	<u>0.27</u>
4-5	51	11	0.22	0.06
*4-6	51	25	0.49	<u>0.67</u>
*4-7	51	33	0.65	<u>0.40</u>
4-8	51	44	0.86	<u>0.33</u>
*5-1	51	36	0.71	<u>0.46</u>
**5-2	51	39	0.76	<u>0.33</u>
**5-3	51	21	0.41	<u>0.27</u>
**5-4	51	34	0.67	<u>0.33</u>
*5-5	51	31	0.61	<u>0.46</u>
*5-6	51	33	0.65	<u>0.60</u>
5-7	51	41	0.80	<u>0.53</u>
*5-8	51	22	0.43	<u>0.54</u>

*6-1	51	17	0.33	<u>0.40</u>
**6-2	51	21	0.41	<u>0.27</u>
*6-3	51	17	0.33	<u>0.40</u>
6-4	51	17	0.33	0.20
6-5	51	24	0.47	0.20
*6-6	51	29	0.50	<u>0.53</u>
6-7	51	26	0.51	0.20
*6-8	51	24	0.47	<u>0.60</u>

表5 文法の項目分析結果一覧表

問題番号	受験者数	正答者数	正答率	識別力
*7-1	51	22	0.43	<u>0.47</u>
*7-2	51	30	0.59	<u>0.60</u>
*7-3	51	23	0.45	<u>0.67</u>
*7-4	51	20	0.39	<u>0.60</u>
7-5	51	23	0.45	0.20
*7-6	51	39	0.76	<u>0.47</u>
*7-7	51	24	0.47	<u>0.40</u>
*7-8	51	20	0.39	<u>0.67</u>
7-9	51	6	0.12	-0.1
**7-10	51	14	0.27	<u>0.27</u>
*7-11	51	27	0.53	<u>0.53</u>
*7-12	51	30	0.59	<u>0.87</u>
*7-13	51	15	0.29	<u>0.40</u>
7-14	51	20	0.39	-0.1
**7-15	51	31	0.61	<u>0.33</u>
8-1	51	32	0.63	0.20
*8-2	51	39	0.76	<u>0.47</u>
*8-3	51	30	0.59	<u>0.53</u>
**8-4	51	14	0.27	<u>0.27</u>
**8-5	51	35	0.69	<u>0.34</u>
8-6	51	25	0.49	0.13
*8-7	51	25	0.49	<u>0.46</u>
*8-8	51	27	0.53	<u>0.47</u>
8-9	51	15	0.29	0.20

*8-10	51	23	0.45	<u>0.53</u>
**8-11	51	30	0.59	<u>0.34</u>
*8-12	51	17	0.33	<u>0.40</u>

項目分析において数値的に望ましい問題項目の目安に合致した問題を、神村(2010)の優良問題と良問題の選別方法³⁾を参考に、優良問題と良問題とに分けた。その結果、文字の問題では優良問題が2問、良問題が1問の計3問(全16問中)、語彙の問題では優良問題が11問、良問題が5問の計16問(全32問中)、文法の問題では優良問題が16問、良問題が5問の計21問(全27問中)であった。

4 まとめと今後の課題

本稿では、今回の作成および実施した試行テストについて述べた。現在までの分析結果としては、文字語彙については、N2レベルの文字語彙では、中国の学生が多かった影響からか正答率が0.8以上の問題が多かった。特に、文字に関しては望ましい目安とされる問題が少なかった。結果として、プレースメントテストとしては、文字に関する試験は不要であると示唆される。

また、今後は、語彙および文法の問題についても、望ましい目安とされる問題数がプレースメントテストで採用されるには足りないため、引き続き、第2回の試行テストを作成し、実施する必要がある。そして、今回はN2レベルとしたが、中国の学生も増えると予想される現在、N2レベルのままが良いのか、また、文字語彙ではなく、今回実施しなかった読解問題も実施する必要があるのかの検討もしなければならない。さらに、ここ数年はベトナムやネパールといった非漢字圏からの留学生が増え漢字を苦手とした学生が多かったことから、1年次科目「日本語特講Ⅰ・Ⅱ」では全クラスで漢字(文字語彙)の学習をすることになっているが、今後、中国の学生が多くなれば、内容を検討することも考えなければならない。

なるだろう。

紀要』 37,77-83

今後は、Web プレースメントテストを開発するために必要なデータ収集を行い、検討を重ねることで、プレースメントテストの妥当性や信頼性を高めるよう努めたい。

<付記> 本稿は、留学生教育学会研究大会（2021年8月21日）において口頭発表した内容に、加筆、修正したものです。

[註]

註1 項目応答理論（IRT: item response theory）は、本テスト作成前に統計的に十分な数の受験者を集めて予備テスト（試行テスト）を行い、テストの結果をもとに、テスト項目ごとの項目困難度（難易度）や項目弁別力（識別力）、当て推量などのパラメータを測定して得点等化を行う。こうして得られた尺度点は、能力が同じであればテストの難易度に関係なく、同じになると言われている。

[引用文献]

- 1) 伊東祐郎（2008）『日本語教師のためのテスト作成マニュアル』アルク,14
- 2) 横山紀子ほか（2011）『学習を評価する』（『国際交流基金日本語教授法シリーズ』12）ひつじ書房,95
- 3) 神村初美ほか（2010）「TMU聴解テスト開発について」『人文学報』428,40-53

[参考文献]

- 1) 和泉元千春ほか（2003）「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—」『日本語国際センター紀要』13,83-93
- 2) 藤田恵ほか（2017）「Webによる日本語プレースメントテストの開発—外国人留学生の受け入れ拡大に向けて—」『立教大学ランゲージセンター