

## 高校生程度の日本語能力テストの開発

——語彙理解テスト・漢字読み取りテストの尺度化——

平 直 樹<sup>1</sup> 小 野 博<sup>1</sup> 前 川 眞 一<sup>1</sup> 林 部 英 雄<sup>2</sup> 米 山 千 佳 子

### DEVELOPMENT OF JAPANESE LANGUAGE TESTS FOR SENIOR HIGH SCHOOLERS' LEVEL

——Calibration of Vocabulary Comprehension Test and Chinese Character Reading Test——

Naoki TAIRA, Hiroshi ONO, Shin-ichi MAYEKAWA, Hideo HAYASHIBE AND Chikako YONEYAMA

The purpose of this study was to construct two Japanese language ability scales, namely, vocabulary and Chinese character recognition, at senior high school level. Four tests were constructed from the original items and administered to about 27,000 Japanese high school students for each scale. The statistical item analysis showed that 242 vocabulary items and 413 Chinese character items could be used for the item response theoretic analysis. IRT analysis of the item responses was carried out separately for each test and the estimated item parameters were equated utilizing the anchor items included in all the tests.

Key words : Japanese language ability, vocabulary, Chinese character, senior high school level, Item Response Theory.

#### I 問題と目的

ことばの能力の測定は、それ自体が教育や研究のテーマと成り得るものである。また、教育に関わる研究を行うとき、測定値が被験者の特性のひとつとして重要な情報であることが多い。それは、ことばの能力が他の様々な能力に密接に関連があると考えられるからである。例えば、日本における学校教育は通常日本語で行われているが、教育される内容を習得するためには、その前提としてそこで使われる程度の日本語を理解し、使用することができなければならないことは自明である。そこで、本研究では、高校生程度の日本語能力を簡便かつ精密に測定することができ、様々な用途で手軽に利用することが可能な尺度を開発した。

ことばの能力という概念は幅広く捉えられるもので

ある。しかし、総合的な能力をそのまま測定するのは極めて困難であるため、本研究では紙筆テスト等の形式で簡単に実施できる基本的な要素に着目した。基本的な日本語能力の中でも、助詞、文型、指示語などの能力は小学校を終える段階でほぼ完成すると考えられる(小野他, 1989)。また、文章作成能力などについては方法論的な問題が多く、簡便な測定尺度を作成することが困難である(平・江上, 1992)。したがって、本研究では、高校生程度の日本語能力を「語彙理解力」と「漢字の読み取り能力」に代表させ、それぞれの尺度を作成することとした。もちろん、「漢字の能力」に焦点を当てて考えた場合、「読み」の能力以外に「書く」能力も重要である。しかし、将来、いわゆる紙筆テストを越えてパソコン等の利用を考えた場合、特に採点の面で大きな困難が生じると考えられる。したがって、漢字に関しては読みの能力を特に取り上げた訳である。

語彙理解テストの項目は、全て五枝選択形式である。幹に示されたことばに最も近い意味の選択枝を選ぶ項

<sup>1</sup> 大学入試センター (The National Center for University Entrance Examinations)

<sup>2</sup> 横浜国立大学 (Yokohama National University)

目を中心に、一部に穴埋め項目を含む。漢字読み取りテストは、問題文の中の一部の漢字にふりがなを振る形式で項目を作成した。語彙理解テスト、漢字読み取りテストの項目の例は、それぞれ FIGURE 1, FIGURE 2 に示す通りである。

ことばの調査

学年組 番号 ( 年 組 番 )  
 氏 名 ( )  
 性 別 ( 男 女 )  
 生年月日 (昭和 年 月 日)

次のことばに一番意味の似ているものを、①②③④⑤の中から1つだけ選んで、マークシートの回答欄にマークしなさい。意味の似ているものが1つより多くある時は、一番良いと思ったもの1つだけにマークして下さい。問題数は全部で73問です。問題番号と回答番号がずれないようによく注意して下さい。

- 1. まどろむ ① ねぼける ② 昼寝する ③ 夢を見る  
 ④ うとうと眠る ⑤ ぐっすり眠る
- 2. 厳密な調査 ① 厳しい ② 詳しい ③ 秘密の  
 ④ 正しい ⑤ 手のこんだ
- 3. 拙い ① 下手くそな ② おさない  
 ③ 良く出来た ④ 準備不足の  
 ⑤ 分かりにくい

FIGURE 1 語彙理解テスト項目の例

- |   |  |   |           |
|---|--|---|-----------|
| 10. 汚れた空気を浄化する。<br>9. 小説を執筆する。<br>8. 光沢のある布。<br>7. もう就寝時刻だ。<br>6. 歓迎会を盛大に催す。<br>5. 食品添加物。<br>4. 聴衆を魅了する。<br>3. 聴衆を魅了する。<br>2. 欧州を旅行する。<br>1. 意見が衝突する。 | 11. 危険の知らせを受ける。<br>12. 港を封鎖する。<br>13. 病魔を克服する。<br>14. 議論が暗礁に乗り上げてしまった。<br>15. 優雅な気分。<br>16. 謙虚な人柄。<br>17. 文明の発祥の地。<br>18. 錠剤を飲む。<br>19. 恐喝事件。<br>20. 3日間の謹慎処分。 | 傍線をつけた漢字にふりがなをふって下さい。<br>性 別 ( 男 女 )<br>氏 名 ( )<br>学 年 組 番 号 ( 年 組 番 )<br>生 年 月 日 ( 昭 和 年 月 日 ) | 漢 字 の 調 査 |
|---|--|---|-----------|

FIGURE 2 漢字読み取りテスト項目の例

尺度化については、被験者の特性に合わせて自由に適応形テスト(野口他, 1983; 平他, 1989)を作成することが可能な項目反応理論(Lord & Novick, 1968; Lord, 1980; Hambleton & Swaminathan, 1985)に基づき、2パラメタ・ロジスティックモデルを用いることにした。したがって、項目特性関数、すなわち、尺度特性  $\theta_i$  を持つ被験者  $i$  が項目特性  $a_j, b_j$  を持つ項目  $j$  に正答する確率  $P_j(\theta_i)$  は、

$$P_j(\theta_i) = [1 + \exp(-Da_j(\theta_i - b_j))]^{-1} \quad (1)$$

と表わされることになる。ただし、 $D=1.7$ の定数である。被験者の尺度特性値を推定するために実施された項目がこのモデルに沿って尺度化されている限り、異なる項目を用いて測定された結果を直接比較することが可能となる。なお、本研究とほぼ同じ方法論を用いた語彙理解尺度に芝(1978)がある。

II 尺度の作成

本研究では、1991年と1992年の2度にわたって項目を作成し、調査を行った。第1回調査の目的は全国から大規模にデータを収集して項目プールの中核となる項目を作成することであり、第2回調査の目的は項目を増やして項目プールを完成させることである。第1回調査の項目分析の詳細は、語彙については平他(1992)、漢字については平他(1993)で既に報告を行っている。

1. 項目の作成

第1回調査においては、国語辞典、新聞、小説等を参考に、語彙90項目、漢字140項目を作成した。

第2回調査では第1回調査の結果を踏まえ、全国から10名の高校教師(うち8名は国語担当)の協力を得て項目を作成し、項目検討会を経て、語彙457項目、漢字434項目を作成した。

なお、語彙の項目については、漢字の読み取り能力の影響を排除するために、漢字にふりがなを振った。

2. 予備調査

五枝選択形式を取った語彙理解尺度の項目は、誤答選択枝の作り方によって項目の善し悪しが変わるため、実施する被験者集団に対する困難度や識別力といった項目の性質についての予想が的中しにくい。したがって、語彙については予備調査を行った。

第1回調査については、実施時期は1991年5月、被験者は東京都内にある私立女子高等学校1年~3年生計268名である。

困難度の指標として正答率、識別力の指標として総

得点と項目得点との点双列相関係数を用いて項目分析をした結果、適当と思われる73項目を本調査に採用した。

第2回調査については、第1回調査で作成した項目の中からアンカー項目として20項目と新たに作成した項目60項目、計80項目から成るテストを6種類作成した。実施時期は1992年7月、被験者は全国各地の公立高校6校の生徒である。1校に1種類ずつの問題冊子を送付した。4校は生徒の大半が大学進学を目指す公立普通科高校、1校は一部に職業科(看護科)の生徒を含む公立高校、1校は農業高校である。被験者数は1校につき172名~271名、計1,474名であった。

第1回調査と同様の手続で項目分析を行い、アンカー項目15項目、新項目180項目を本調査に採用した。

### 3. 本調査

第1回調査の実施時期は、1991年7月~10月、被験者は全国各地の28の高校の生徒計18,293名である。漢字についてはそのうち12,486名分を分析の対象とした。28校中3校が私立進学校、21校が生徒の大半が大学進学を目指す公立普通科高校、1校が商業高校、3校が農業高校である。第1回調査では、全ての被験者に同じ項目から成るテストを実施した。項目数は、語彙が73項目、漢字が140項目である。

第2回調査の実施時期は、1992年9月~11月である。被験者は全国各地の18の高校の生徒計8,685名である。18校中1校が私立進学校、7校が生徒の大半が大学進学を目指す公立普通科高校、4校が進学希望者と就職希望者が混在する公立普通科高校、5校が就職希望者が大半の公立普通科高校、1校が農業高校である。第2回調査では、語彙、漢字ともに難しい順にA、B、Cの3種類のテストを作成した。Aを6校、Bを8校、Cを9校に実施した。生徒を2分してA、Bの両方を実施した学校が2校、B、Cを実施した学校が3校あった。それぞれの被験者数はAが2,978名、Bが2,822名、Cが2,870名である。項目数は、語彙は各版とも75項目、漢字はAが125項目、B、Cがそれぞれ105項目である。

なお、語彙の15項目と漢字の20項目が、第1回調査のテストと第2回調査のA、B、Cの各版に共通のアンカー項目であり、被験者と項目の構成はFIGURE 3のようなデザインになっている。

### 4. 項目分析

語彙、漢字項目について、第1回調査と第2回調査のA、B、C各版の平均得点、標準偏差の値はTABLE 1に示す通りである。いずれも全体としては適切な困

	アンカー項目	非共通項目				合計
		第1回	第2回A	第2回B	第2回C	
第1回調査	語彙	15項目	58項目			73項目 = 140項目
	漢字	20項目	120項目			
第2回調査A版	語彙	同上	60項目			75項目 = 125項目
	漢字	同上	105項目			
第2回調査B版	語彙	同上	60項目			75項目 = 105項目
	漢字	同上	85項目			
第2回調査C版	語彙	同上	60項目			75項目 = 105項目
	漢字	同上	85項目			

FIGURE 3 本研究のデザイン

TABLE 1 平均得点・標準偏差

	第1回調査	第2回調査 A	第2回調査 B	第2回調査 C
語彙項目数	73	75	75	75
語彙平均得点 (正答率%)	41.00 (56.16)	42.13 (56.17)	42.03 (56.04)	51.30 (68.40)
語彙標準偏差 (SD%)	10.52 (14.41)	13.26 (17.68)	14.21 (18.94)	14.25 (19.00)
漢字項目数	140	125	105	105
漢字平均得点 (正答率%)	88.43 (63.16)	61.85 (49.48)	54.74 (52.13)	57.28 (54.55)
漢字標準偏差 (SD%)	26.76 (19.11)	26.12 (20.90)	22.89 (21.80)	25.00 (23.81)

難度の項目群であった。

本調査の項目分析においては、予備調査と同様の分析に加え、設問解答率分析図(越田他, 印刷中)を用いた。設問解答率分析図とは被験者集団を総得点によって5段階に分類し、各項目の各段階における各選択枝の選択率(語彙)、またはその項目に対する正答率(漢字)を折れ線グラフで図示して比較するものである。それによって被験者全体に対する識別力のみならず、それぞれの部分集団に対する項目の識別力を知ることができる。すなわち、正答選択枝の場合、上位群になるにしたがって選択率が高く、誤答選択枝はその逆であることが望ましい。例えば、FIGURE 4は語彙理解尺度のアンカー項目として実施された項目であるが、正答選択枝の選択率から、A版においては最下位群と下位群との識別力が最も高く、B版においてはほぼ全体的、C版においては上位群と最上位群との識別力が最も高い性質を持つ項目であることがわかる。また、誤答選

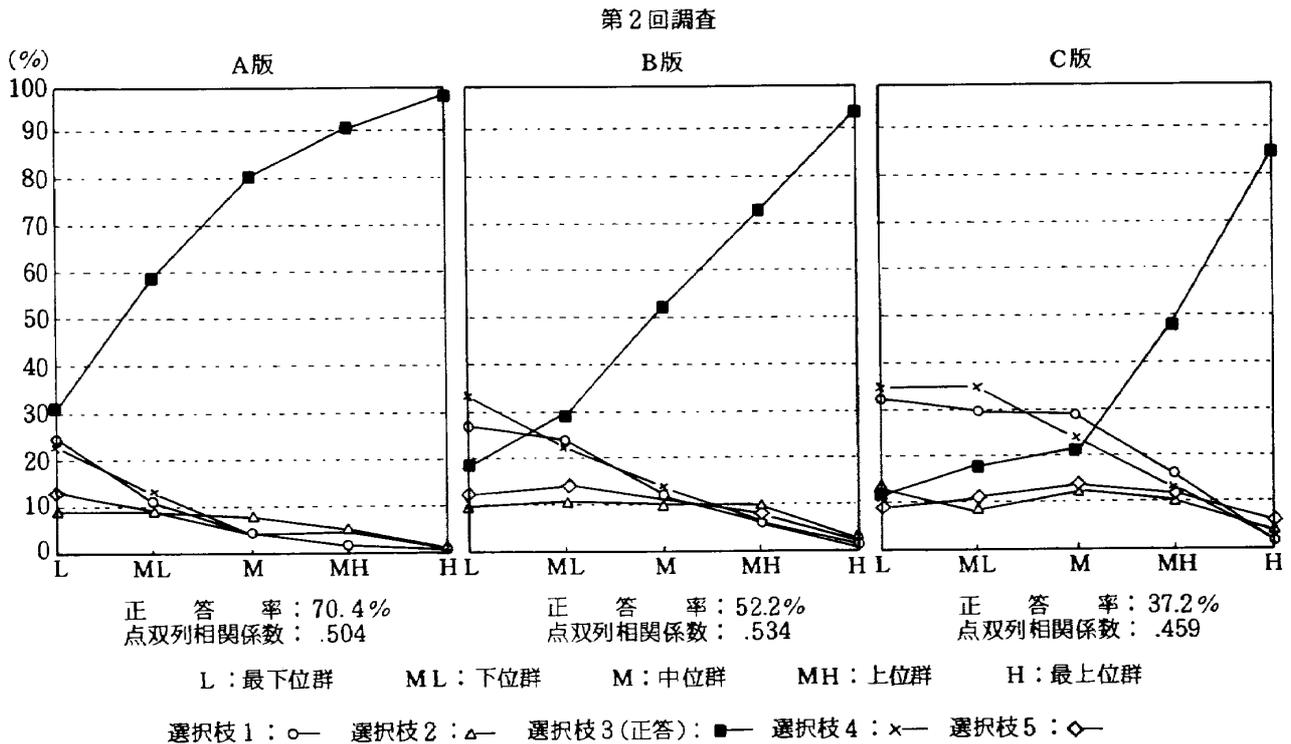


FIGURE 4 設問解答率分析図の例 (語彙アンカー項目「ねぎらう」の場合)

択枝についても選択枝1と4が最もよく機能していたことが分かる。

これらの分析の結果、識別力が低かった項目、内容的に問題があった項目について、語彙11項目、漢字2項目を項目プールから除外することとした。その結果、最終的に出来上がった尺度化のための項目プールは、語彙242項目、漢字413項目となった。

### 5. 尺度化と等化

本研究の実施デザインは、FIGURE 3で示したように、4群の被験者に対して実施した項目が一部重なる共通項目デザイン(野口, 1991)となっている。本研究では項目、被験者とも数が多いため、項目パラメタを全て一度に推定することは不可能である。したがって、最初に各項目・被験者群ごとに項目パラメタを推定し、その後、4つの尺度を等化する手続が必要となる。項目パラメタの推定には周辺最尤推定法によるプログラム BILOG (Mislevy & Bock, 1989) を用い、尺度の等化には前川・鈴木の方法(前川・鈴木, 1988; 前川, 1991)を用いた。被験者の母集団分布は、語彙、漢字ともに平均0、分散1となるように固定した。なお、第1次調査のデータに関しては、語彙、漢字とも3,000名分のデータを無作為抽出して用いた。

項目と推定されたパラメタのリストの一部を、語彙の項目は TABLE 2、漢字の項目は TABLE 3 に示す。

### III まとめと課題

本研究の被験者は厳密な無作為抽出の手続によって全国の高校生の母集団から得られたものではないが、ほぼ標準的な高校生の集団を網羅していると考えられる。したがって、本研究で開発した尺度は、高校生程度の日本語能力の基礎を簡便に測定するという目的で用いられるならば、十分に実用であると考えられる。測定対象となる被験者の性質に応じ、パソコン等の機器を利用した適応形テストを作成することも可能である。

今後の課題としては、発達の関心に対応できるように、中学生以下の子どもの能力を測定する項目も付け加えた、より広い範囲をカバーする項目プールを作成することが必要だと思われる。また、本研究で開発した尺度で測定される能力と、作文能力など言語能力の他の側面との関連について調査することも重要である。

TABLE 2 語彙理解尺度項目プール (抜粋)

項目	a	b
1 優柔不断	1.13	-2.38
2 お目玉を食う	1.11	-2.10
3 そうは問屋がおろさない	0.28	-1.88
4 焼け石に水	1.31	-1.83
5 携わる	1.16	-1.76
6 助言	0.57	-1.53
7 生き字引	1.32	-1.47
8 はにかむ	0.93	-1.42
9 さげすむ	0.98	-1.35
10 矢継ぎ早の	0.84	-1.27
11 匙を投げる	1.10	-1.23
12 すこぶる	0.84	-1.22
13 はかない	0.81	-1.10
14 露骨に	0.99	-1.05
15 泡を食う	0.41	-1.02
16 おびただしい	0.91	-0.97
17 腰が低い	0.97	-0.91
18 無い袖は振れぬ	0.63	-0.89
19 ねずみ算	0.75	-0.83
20 理不尽	0.95	-0.78
21 逆鱗に触れる	0.54	-0.77
22 さばをよむ	0.48	-0.68
23 巧拙	0.56	-0.65
24 喉元を過ぎれば熱さを忘れる	0.83	-0.57
25 青菜に塩	0.59	-0.53
26 鶴の一声	0.74	-0.50
27 辟易する	0.62	-0.45
28 一蹴する	0.68	-0.39
29 したためる	0.62	-0.36
30 体裁	0.58	-0.31
31 虻蜂とらず	0.32	-0.24
32 とらぬ狸の皮算用	0.68	-0.20
33 妙案	0.83	-0.13
34 時化	0.65	0.02
35 嗚咽	0.91	0.08
36 矢の催促	0.66	0.11
37 邪険	0.76	0.34
38 辛辣	0.71	0.40
39 及第	0.83	0.55
40 毫も	0.61	0.67
41 人を呪わば穴二つ	0.35	0.75
42 未曾有	0.83	0.92
43 ジレンマ	0.70	1.01
44 詭弁	0.45	1.12
45 こころもとない	0.45	1.17
46 遜色ない	0.30	1.39
47 子はかすがい	0.50	1.59
48 尊大	1.15	1.61
49 言わぬが花	0.48	1.84
50 実るほど頭を垂れる稲穂かな	0.36	2.21

TABLE 3 漢字読み取り尺度項目プール (抜粋)

項目	a	b
1 封鎖	0.88	-3.05
2 光沢	1.23	-2.72
3 恐喝	1.10	-2.47
4 凝視	0.62	-2.20
5 発作	1.45	-1.95
6 危篤	1.08	-1.77
7 滑らか	0.82	-1.68
8 土壇場	1.14	-1.56
9 会得	1.11	-1.38
10 浸る	1.19	-1.30
11 迅速	1.12	-1.10
12 溢れる	0.78	-1.02
13 怪我	1.43	-0.90
14 緩和	0.97	-0.81
15 承る	0.58	-0.76
16 時雨	0.94	-0.70
17 貴賓	1.23	-0.60
18 転嫁	1.21	-0.55
19 拘留	1.00	-0.49
20 彩り	0.94	-0.44
21 不憫	1.29	-0.36
22 担う	1.30	-0.33
23 流浪	1.21	-0.31
24 索性	0.97	-0.24
25 寡黙	1.19	-0.22
26 玄人	0.64	-0.13
27 崇める	0.84	-0.09
28 厳か	0.81	-0.02
29 忽然	1.39	0.00
30 対峙	1.08	0.07
31 驚愕	1.49	0.12
32 無精	1.39	0.14
33 踏襲	0.81	0.26
34 拮抗	0.96	0.27
35 佇む	1.13	0.33
36 謁見	1.63	0.36
37 廉価	1.00	0.46
38 慟哭	1.33	0.50
39 猜疑	1.39	0.64
40 督促	0.39	0.70
41 諮問	0.95	0.77
42 破綻	1.23	0.90
43 閉塞	1.03	0.95
44 弔問	1.08	1.08
45 頹廃	1.54	1.29
46 敬虔	1.24	1.39
47 蘊蓄	1.33	1.50
48 奢る	1.24	1.66
49 反駁	1.21	1.83
50 敷衍	1.32	2.22

## 参考文献

- Hambleton, R.K., & Swaminathan, H. 1985 Item Response Theory —Principles and Application —, Kluwer Nijhoff Publishing.
- 越田 豊・池田輝政・山田文康 印刷中 平成元～3年度の共通一次・センター試験「生物」試験問題の性能分析 生物教育, **34**, No.1 (1993—1994).
- Lord, F.M. 1980 Applications of Item Response Theory to Practical Testing Problems., Hillsdale, NJ : Lawrence Erlbaum Associates Inc..
- Lord, F.M. & Novick, M.R. 1968 Statistical Theories of Mental Test Scores., Reading, Massachusetts : Addison-Welsey.
- 前川真一 1991 パラメタの推定 芝祐順編 項目反応理論 基礎と応用 東京大学出版会 87—129.
- 前川真一・鈴木規夫 1988 独立に推定されたテスト項目パラメタの共通尺度上での等化法 —共通第1次学力試験の選択科目の差異の分析への適用— 大学入試センター研究紀要, **17**, 249—271.
- Mislevy, R.J. & Bock, R.D. 1989 PC-BILOG3 : Item analysis and test scoring with binary logistic models. Mooresville, ID : Scientific Software.
- 野口裕之 1991 項目反応理論に基づくテストの作成法 芝祐順編 項目反応理論 基礎と応用 東京大学出版会 51—86.
- 野口裕之・芝 祐順・丹 直利 1983 語彙理解尺度の研究II —項目固定版と適応形テストによる測定— 東京学芸大学紀要, **34**, 101—114.
- 小野 博・繁樹算男・林部英雄・岡崎 勉・市川雅教・木下ひさし・牧野泰美 1989 日本語力検査の開発 昭和61—63年度科学研究費報告書試験研究(1) 課題番号 6181003 (未公刊)
- 芝 祐順 1978 語彙理解尺度作成の試み 東京大学

教育学部紀要, **17**, 47—58.

- 平 直樹・江上由実子 1992 ESSAY TEST の方法的諸問題に関する研究の動向について 教育心理学研究, **40**, 108—117.
- 平 直樹・武井澄江・荻野美佐子 1989 適応形質問紙形式の乳幼児用言語理解尺度への応用 教育心理学研究, **37**, 392—399.
- 平 直樹・小野 博・林部英雄 1992 高校生用日本語語彙理解力テストの開発 —(1)試作問題の精選— 大学入試センター研究紀要, **21**, 107—143.
- 平 直樹・小野 博・林部英雄 1993 高校生用日本語能力テストの開発 —(2)漢字読み取りテストの項目分析と誤答パターンについて— 大学入試センター研究紀要, **22**, 53—85.

## 付 記

ページ数に制約があるため、実際にテストを作成し、実施するための情報を全て掲載することはできなかった。より詳しい資料は、大学入試センター研究開発部小野博宛に請求していただきたい。

## 謝 辞

本研究を進めるにあたり、全国の多くの高校の先生方と生徒のご協力を得ることができました。感謝いたします。とりわけ、第2回調査の項目作成に協力してくださいました、浅井裕先生、上原巖先生、尾形昇先生、佐澤美彦先生、多田久実子先生、永田栄一先生、福津範人先生、藤原慶久先生、森義敬先生、屋比久貞夫先生(五十音順)に感謝致します。最後に、問題冊子の発送やデータの整理、入力等に協力して下さった、富田治子さんはじめ大学入試センター研究開発部の非常勤職員の方々、横浜国立大学の学生の皆様に感謝したいと思います。

(1994.4.11受稿, 9.17受理)