台灣地區日語學習及使用者之日語能力調查
—FLPT 研究計畫成果報告—

藍碧珠
財團法人語言訓練測驗中心 編審

中文摘要

本研究旨在了解台灣地區日語學習及使用者的聽力、用法、字彙閱讀三項技能之發展狀況，建立以台灣成人日語學習及使用者爲母體之語言能力常模。研究計畫對象包含日語主修及非主修之各大學學生及公民營機構在職人士。

本計畫使用的評量工具爲標準化常模參照測驗 FLPT 日語能力測驗，抽樣方式採叢聚抽樣，受試者包含台灣北、中、南、東部各大學日語學習者、公民營機構日語使用者，有效樣本 297 名。

本研究於民國九十四年完成資料收集及施測，測驗結果顯示，日語全體受試者以聽力測驗成績最佳，其次依序爲字彙閱讀、用法。平均成績分別爲 65 分、58 分、57 分。在學者及在職者之子群組中，前者以大學三年級、後者以學習時數「400~450 小時」組的成績最接近全體日語組受試者的平均成績。其他各項背景之成績交叉分析與發現，詳載於本文。

關鍵字：FLPT、常模參照測驗、叢聚抽樣、日語學習者、日語使用者

Investigating the Language Development of Japanese Learners and Users in Taiwan Using the Foreign Language Proficiency Test

P.C. Lan
Editor, The Language Training and Testing Center

Abstract

The purposes of this study were to investigate the language development of Japanese learners and users in Taiwan, especially in the three skills of listening, usage, and vocabulary and reading (VR), as well as to establish a norm based on the performance of Japanese learners and users in Taiwan. The population of interest includes university students, including Japanese majors and non-Japanese majors, and employees in both the public and private sectors.

The Foreign Language Proficiency Test (FLPT), a standardized norm-referenced test, was used as an instrument in this study. The cluster sampling approach was utilized to select 297 Japanese learners and users to participate.

Data collection was completed in 2005. The results show that the participants performed best on the listening test, while their performances on the usage test were the second highest, and that on the VR test, third. The average scores of the three tests were 65, 58, and 57, respectively. The subjects were divided into "student" group and "employee" group and further analyses were conducted. It was found that the performance of the subgroup of "juniors" of the former, and that of the subgroup of "employees who have learned Japanese for 400-450 hours" of the latter were closest to the overall performance of all subjects. Detailed analyses and findings could be found in the paper.

Keyword: Foreign Language Proficiency Test (FLPT), norm-referenced test, cluster sampling, Japanese learners, Japanese users

台湾における日本語学習・使用者の語学力調査
―FLPT 研究計画の結果報告―

藍　碧珠
財団法人語言訓練測験中心　編審

要　旨

台湾における日本語学習・使用者の聴力、用法、語彙・読解での能力分布がどうなっているかを調査し、台湾の日本語成人学習・使用者を母集団とする言語能力基準を構築するため、2005 年に語言訓練測験中心より日本語専攻と非日本語専攻の学習・使用者を含めた各大学と公民営機構の協力を得て、FLPT 二期研究計画が実施された。

評価のツールとして、標準化された団体準拠基準テストである FLPT 外国語能力試験が使われ、サンプリングは、多段サンプリング調査法が採用された。調査対象として、台湾の北部、中部、南部、東部にある各大学、公民営機構での日本語学習・使用者から代表的な受験者が合わせて 297 名抽出された。

2005 年に二期研究計画のデータ収集とテスト実施を行なった。テストの結果、全受験者を通して、聴力の成績が最も良く、語彙・読解と用法の成績がそれに続いている。平均点はそれぞれ 65 点、58 点、57 点である。在学者と在職者の下位グループのうち、全受験者の平均得点に最も近接しているのは、それぞれ大学 3 年生と学習時間数「400~450 時間」グループとなる。その他のバックグランドによる成績分析と発見については、本文を参照されたい。

キーワード：FLPT、団体準拠基準テスト、多段サンプリング調査、
日本語学習者、日本語使用者

台湾における日本語学習・使用者の語学力調査
—FLPT 研究計画の結果報告—

藍　碧珠
財団法人語言訓練測験中心　編審

1.　研究の動機と目的

台湾における日本語学習・使用者の聴力、用法、語彙・読解での能力分布がどうなっているかを調査し、台湾の日本語成人学習者を母集団とする言語能力基準を構築するため、2005 年に語言訓練測験中心より日本語専攻と非日本語専攻の学習・使用者を含めた各大学と公民営機構の協力を得て、FLPT二期研究計画[1]が実施された。

この研究計画の実施により、第二外国語学習の促進につながり、当方で開発した「外国語能力テスト（FLPT）」能力尺度の考察に活用され、台湾における外国語学習者がより利用しやすい外国語能力の検定試験の作成に役立つかと考える。

2.　研究方法

2.1　評価ツール

二期研究計画は一期研究計画と同じく、客観的な評価ツール「FLPT」外国語能力テストを採用した。FLPTは大型の標準化[2]された集団基準準拠テスト[3]（standardized, norm-referenced test）で、テス

---

[1] 2004 年に実施された「台湾の高等教育機関における欧日言語専攻者を対象とする語学力評価」研究計画に引き続き、今回の研究計画が二期として実施された。

[2] Borg & Gall によると、標準化されたテストの特徴として、客観性、試験実施条件の一致、そしてパーセンタイル順位によって作成した基準及び信頼性、妥当性を有することである。（Borg & Gall、1989）

[3] 点数を解釈する方法によって、テストは二種類に分類される。一つは、集団基準準拠テストで、このテストの類型として、TOEFL がよく挙げられる。一つは、目標基準準拠テスト（criterion-referenced test）で、全民英検がこの類に属する。集団基準準拠テストにおける受験者の得点は、素点ではなく、集団基準に準拠した上で、変換された標準得点である。その標

トする言語に、英、日、仏、独、スペイン語の五つがあり、広く公民営機関に採用されている言語能力検定ツール[4]で、国内で満 18 歳の大学程度の外国語学習者の能力検定に見合った試験内容となっている。

FLPT の聴力、用法及び語彙・読解という三項目の筆記テストは、マークシート方式で、コンピューターで採点され、各テストで独立採点し、「標準得点」で計算される。平均得点が 60 点で、20 点を標準偏差とする。各項目の得点スケールは、0~120 点とする。

FLPT聴力テストは、日常生活で必要とされる一般的な[5]外国語聴解能力が備わっているかどうかをテストするものである。ある単文の発話に対する返答力、単独の発話が伝えようとしている意味、そして短い会話や一連の発話の主旨や細部情報などへの理解力をテストすることを目的とする。

用法テストは、個別項目[6]問題（discrete-point item）形式で、一般的によく使われている文法や文型への理解力をテストすることを目的とする。

語彙・読解テストは、語彙と読解の二部からなっている。語彙力テ

準得点をもって、受験者のこの基準の中での相対的な位置を表わす（王文科、1999）。

[4] 現在、400 余の公民営機構が FLPT を採用している。教育部の公費留学試験、台湾大学教育プログラム、外交部、衛生署、司法院、台湾電力会社、中華電信、中国原油及び 10 余の銀行などが採用している。考選部公務人員保訓会で、各級公務機関、学校による公務員の海外学習派遣選考の際の言語能力証明として指定されている。

[5] General language proficiency は specific language proficiency に対する相対的な概念である。前者のテストはある特定した分野の専門知識を組み入れない。たとえば、FLPT や全民英検がそうである。後者は特定した分野の専門知識をテスト問題に組み入れる必要がある。例えば、添乗員外国語テストと GMAT 等である。

[6] 個別項目問題（discrete-point item）は、テンスや品詞などのような単一の特定した言語項目についてテストするものである。その利点として、当該言語からサンプリングできるスポットが広く、受験者が個別項目での解答成績は、他の問題への表現に影響を与えて二重減点（double penalty）を引き起こすことはない、などが挙げられる。

ストは個別項目問題で、読解力テストは、前後の文脈から語彙の意味を読み取る力､及びリーディングスキルの駆使により文章の主旨、あらましや細部情報を把握できるかどうかをテストすることを目的とする。

筆記テストの問題様式や問題数などについては表 2.1 を参照されたい。

表 2.1　FLPT 筆記テスト問題様式と問題数

テスト項目	問題様式	問題数	テスト時間
聴力	I.　応答	15	約 30 分
	II.　理解	15	
	III.　会話及び長文	20	
用　法	文法､文型パターン	80	40 分
語彙・読解	I.　語彙：単語､慣用語句	40	60 分
	II.　読解：6 篇の文章	20	

2.2　サンプリング

二期研究計画のサンプリング対象として、在学の日本語専攻者のほか、在学の非日本語専攻者及び在職者が含まれた。

サンプリング方法として、多段サンプリング調査（クラスターサンプリング）が採用された。即ち、各学校、機構を一つのまとまった抽出グループとして、各言語、機構から調査対象がそれぞれ 5 名推薦された。これにより、各機構、団体の行政システムを通して始めてバランスの取れた、且つ代表性のあるサンプルの抽出が可能となるのである（王文科、1999）。

研究開発チームから欧日言語を統合した研究計画[7]において 450 の学校及び公・民営企業に調査依頼を出した結果、日本語グループについて 110 の団体から協力との回答を得た。さらに異な

[7] 二期研究計画に、日本語のほか、フランス語、ドイツ語、スペイン語が含まれているため、サンプリング作業は一括して行われた。

る領域における在学者と在職者の差異を分析するために、その下にそれぞれ二つの部分群に細分する。各部分群の定義及び参与学校・機関数につき表 2.2 を参照されたい。

表 2.2　各部分群の参与機構数

在学-日本語専攻学科数	在学-非日本語専攻学科数*	在職-公営機関数	在職-民営企業数	団体合計数
22 (20%)	15 (14%)	69 (63%)	4 (4%)	110

*第二外国語課程、語言センター等

在学者のサンプリング対象の抽出は、協力大学の教師から、2 年生以上の学生を推薦してもらうこととした。在職者については、協力機構から、専攻分野を問わず、職場で日本語を 250 時間以上学習している勤続者を募集することとした。[8]

2.3　下位群の定義

専攻分野と職業団体の違いによる日本語能力を分析するため、すべての受験者を日本語専攻かどうかによって区分するほか、学習年数によって在学者を大学 2 年から 4 年の 3 グループに、在職者を学習時間によって 5 グループに分ける。各下位群のうち、学制の違う各学年をまとめるため、在学者の具体的な定義については表 2.3 を

[8] 一期計画の結果から、FLPT の受験に向いている者は真性初心者（true beginners）ではなく擬似初心者（false true beginners）だということが分かった。後者の学習期間が長いが、実際にその言語を使ったことがなかったり、長年触れていなかったりしているため、パフォーマンス力は初心者と似ている。前者は、学習期間が短く、ごく基本的な語彙または文法概念しかない。台湾での例をあげて言うと、教育政策と課程の規定により、欧日語の学習者は、ほとんど大学時期から当該言語に初めてアプローチしたため、大学 1 年は、真性初心者と言える。英語学習者の場合は少々異なる。ほとんど小、中学から英語教育を受け、学習年数は長いが、大学になっても英語での基本的な応対、コミュニケーションが取れない場合は、擬似初心者と言える。

参照されたい。

表 2.3　在学者-学習年数による分類

コード	言語能力	学制と学年
02	大学 2 年に相当	大学 2 年、四技 2 年、五専 2・3 年
03	大学 3 年に相当	大学 3 年、四技 3 年、二技 1 年、五専 4・5 年
04	大学 4 年に相当	大学 4 年、四技 4 年、二技 2 年

2.4　テスト実施のプロセス及び変動要因のコントロール

研究計画で集めたデータの代表性を確保し、母集団の特性を適切に反映することができるように、サンプリング方法と評価ツールの選別に配慮する他、テストの実施について、一期計画の経験から、非言語関連の要因に伴う受験者のパフォーマンスへの影響を最小限にとどめるために、下記いくつかの変動要因を調節した。

1）動機要因（motivation effect）

一期研究結果で、「受験動機」は受験者のパフォーマンスへの影響が大きいことが分かった。動機が強いほどパフォーマンスが良く、逆に動機が弱いとパフォーマンスも悪い（語言訓練測験中心、2004）。そのため、二期計画のサンプリングにおいては、各部分群や個人の動機の差異による成績誤差をなくすために、受験者の自主的な出願という形に切り替えた。

また、実験的な性質が強い[9]一期計画と違い、二期計画では、受験者を正規のFLPTテストに編入した。後者では、テスト結果の成績が正式の効力を有するために、受験者の「受験動機」が高くなり、全

[9] 即ち、正規の FLPT の実施プロセスと完全に一致するものではなく、一期計画のテストの実施プロセスは、研究目的に沿ってデザインされたものである。テストの結果は研究のためだけに使用が限定され、受験者の成績、背景資料なども関係のない第三者への公開は全くない。受験者は学校の実施担当を通して自分の成績を知ることはできるが、それをもって、語言訓練測験中心に正式の効力のある FLPT 成績証明を申請することはできない。

力を出して解答することで、結果がより実質的に受験者の能力を反映することができると考えられる。

2）時間と環境要因（effect of time and site）

サンプリングテスト実施時期について、一期計画と同じように、各大学の中間テスト終了後の勉学効果ピーク時期の 4、5 月に行われることにした。この時期の学生の能力は、学年終了時のレベル（exit level）に相当している。また、中間テスト或いは期末テストの前に実施する場合、参加意欲が低下したり、卒業が間近で参加できなくなったりするというようなことも回避できた。

受験者は 94 年 4 月 16 日または 5 月 7 日のいずれかのFLPTテストを選ぶことができる。ただし、同一言語で且つ同一の団体で出願した場合は、同一の期日の受験に限る。すべてのサンプリングテストは、語言訓練測験中心で実施された。このことにより、一期計画で各大学でそれぞれ実施した回数延べ 16 回、実施期間延べ 5 週間に比べると、今回のほうが遥かに期間短縮[10]となると同時に、実施環境による変動幅も小さくなった[11]。

3）疲労要因（fatigue effect）

一期計画でのテスト時間は、約 3 時間[12]かかった。その中で、授業日の午後でのテスト実施において受験者がかなり疲労していたことで、第 3 セクションの語彙・読解テストのパフォーマンスに影響が出たということが分かった。そのため、二期計画では、各項目テストの間に行われる予定のアンケートを取り消した代わりに、出願

---

[10] サンプリング期間が短縮した後、学習要因（growth effect）も安定してきた。サンプリング期間が長引いたことや、学友、同僚間の意見交換などによる各部分群間の学習差異によってテスト後の対照比較ベースが異なるという現象が回避できた。

[11] 一期計画では、各大学、学科の授業に合わせて、試験会場の手配などもあることで、環境要因が一致しないことが多かった。たとえば、一般の教室や、ラボ教室などの違い、実施時間についても、午前だったり、午後だったりして、或いは授業日と週末の違い、晴天と雨天の天候の差異など、情況が様々で一致しなかった。

[12] 各項目テストの後すぐアンケートの記入があった。

時にアンケート項目の記入を依頼した。これにより、全テスト時間は約 130 分に短縮した。また、テスト実施日は、受験者の授業または仕事などがない土曜にし、疲労要因によるテストパフォーマンスへの影響を最小限にとどめた。

4）テストのプロセスと実施条件の要因（effect of order and administration）

二期計画において、参加者が正規のFLPTテストを受験するため、標準化されたテストのプロセス[13]と手順を踏んで実施された。試験監督も語言訓練測験中心のメンバーが担当した。即ち、各試験会場の実施条件が一致し、聴力放送の音量、テスト時間の把握及び指示用語など、すべて監督員手引きに基づいて行われたため、上述の諸実施条件による受験者パフォーマンスへの影響を最小限にとどめることができた。

2.5　研究課題

台湾における日本語の成人学習者を集団とした言語能力基準を構築するため、二期計画の研究課題は下記の通りである。

1）　台湾における 18 才以上の日本語学習・使用者（真性初心者を含まず）の聴力、用法、語彙・読解という三つの技能の学習、使用情況はどうか。

2）　FLPT はこの集団の能力を検定するのに適しているかどうか。

3）　在学者と在職者の言語能力は異なっているか。どう異なっているか。

4）　職場での使用状況はどうなっているか。

5）　異なるバックグランドの学習者或いは使用者の間で、三項目の技能の発展情況は異なっているか。どう異なっているか。

[13] 即ち、聴力、用法、語彙・読解の三項目テストの順序で休憩を入れずに行われた。

3. 研究結果

3.1 全体的な統計分析

日本語の実際受験者数は297名で、三項目テストを実施した結果、聴力の成績が最も良く、語彙・読解が次に、用法が最後に並んだ。聴力テストにおいて、平均値 64.92 点、標準偏差 24.70 点、中央値[14]と最頻値[15]が同じく 67 点で、尖度[16]-1.07、歪度[17]-0.18 であることは、分布曲線が正規分布よりやや平たんで、やや得点の高い区域に集中していることを示している。最高点と最低点の開きが 101 点であることは、三項目テストの中で受験者能力値の分布範囲が広く、個々の差異が最も大きいことを示している。

用法テストの成績において、平均値 57.20 点、標準偏差 21.66 点、中央値 56 点、最頻値 35 点、尖度-0.82、歪度 0.23 であることは、分布曲線がやや平たんで、得点の低い区域に集中しているものの、正規分布と言える。最高点と最低点の開きが 89 点であることは、三項目テストの中で受験者能力値の分布範囲が最も小さく、個々の差異が最も小さいことを示している。

語彙・読解テストにおいて、平均値 58.30 点、標準偏差 21.73 点、中央値が 58 点で平均値とほぼ一致し、最頻値 66 点、尖度-0.88、歪度 0.03 であることは、分布曲線が正規分布よりやや平たんであることを示している。最高点と最低点の開きが 95 点で、受験者能力値の分布範囲が聴力、用法の両テストの間にあり、個々の差異は聴力より小さいことを示している。表 3.1 と図 3.1 を参照されたい。

表 3.1 各テスト項目の成績分析

---

[14] 中央値（Median）、受験者全体の成績の中で一番中央の位置にある受験者成績のことを言う。（林清山、1999）

[15] 最頻値(Mode)、出現した回数が最も多い数値で、即ち最も多くの受験者が得た点数のことを言う。（林清山、1999）

[16] 尖度(Kurtosis)、分布曲線が正規分布であるかどうかを表わす数値の一つである。0 に近いほどより正規分布となる。

[17] 歪度(Skewness)、分布曲線が正規分布であるかどうかを表わす数値の一つである。0 に近いほどより正規分布となる。

項目	聴力	用法	語彙・読解
平均値	64.92	57.20	58.30
中央値	67	56	58
最頻値	67	35	66
標準偏差	24.70	21.66	21.73
分散	609.91	469.20	472.07
尖度	-1.07	-0.82	-0.88
歪度	-0.18	0.23	0.03
範囲	101	89	94
最低点	7	17	10
最高点	108	106	104
標準誤差	1.43	1.26	1.26
受験者数	297	297	297

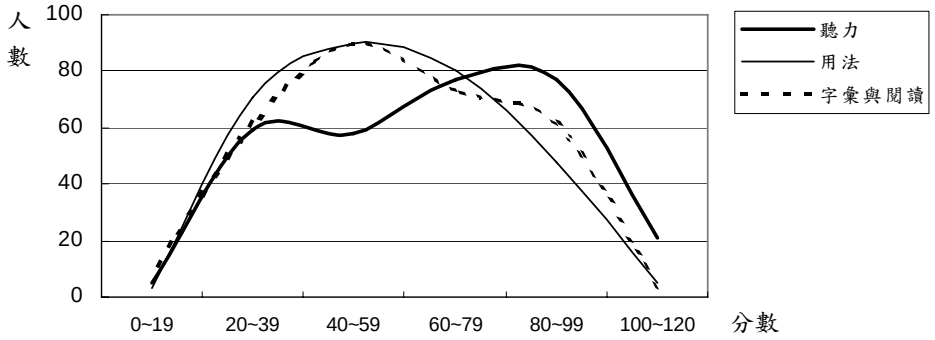

圖 3.1　各テスト成績分布図

教育部公費留学[18]の合格点 60 点を基準として見た場合、聴力、用

[18] 公費留学は 2000 年から、英、日、仏、独、スペイン語の五つの言語の筆記テストを中止した代わりに、教育部が指定した言語検定機構発行の言語能力証明を提出することにした。語言訓練測験中心が即ち教育部の指定した言語検定機構の一つで、FLPT のテスト成績がその指定した合格言語能力証明の一つとなる。公費留学試験の FLPT 合格基準として、聴力、用法、語彙・読解の各テストの平均成績が 60 点以上（総合点数 180 点以上）で、口語成績 S-2 以上とされている。

法、語彙・読解は、それぞれ 175 人、133 人、138 人がこの基準を超えており、それぞれ受験者総数の 59%、45%、47%を占めている。これは、受験者にとって聴力テストが最もやさしく、次が語彙・読解テストで、最も難しいのは用法テストであることを表わしている。各テストの得点分布については表 3.2 を参照されたい。

表 3.2 各テストの得点分布

得点範囲	聴力テスト			用法テスト			語彙・読解テスト		
	人数	人数累計	累計%	人数	人数累計	累計%	人数	人数累計	累計%
110~120	0	0	0%	0	0	0%	0	0	0%
100~109	21	21	7%	5	5	2%	4	4	1%
90~99	35	56	19%	20	25	8%	22	26	9%
80~89	42	98	33%	28	53	18%	39	65	22%
70~79	42	140	47%	30	83	28%	33	98	33%
60~69	35	175	59%	50	133	45%	40	138	47%
50~59	30	205	69%	47	180	61%	51	189	64%
40~49	28	233	79%	43	223	75%	39	228	77%
30~39	36	269	91%	44	267	90%	44	272	92%
20~29	23	292	98%	27	294	99%	17	289	97%
10~19	4	296	100%	3	297	100%	8	297	100%
0~9	1	297	100%	0	297	100%	0	297	100%

3.2 各試験の適用可能性の分析

聴力、用法、語彙・読解の各テストの平均有意確率 P 値[19]がそれぞれ約 0.6 となっており、各テストの難易度が適切であることを表

[19] P-value（有意確率）、試験項目の難易度を表わす数値。数値範囲は 0 から 1 までの間とする。0 に近いほど難しく、1 に近いほど易しい。0.5 に近づくほど、試験項目の難易度が適切だということを示す（余民寧、2002）。Mean P はテストにおける全体の試験項目 P 値の平均値である。

わしている。平均Biserial値[20]が 0.4 以上で、各テストの識別率が良く、受験者能力を有効に識別することを表わしている。各項目テストのAlpha値[21]がそれぞれ 0.90 以上ということは、得点の信頼性が高く、FLPTが受験者の言語能力を推測するのに適していることを表わしている。表 3.3 を参照されたい。

表 3. 3　各試験の適用可能性の分析

統計數値	聴力テスト	用法テスト	語彙・読解テスト
Mean P-value	0.642	0.570	0.603
Mean Biserial	0.463	0.395	0.403
Cronbach's Alpha	0.904	0.918	0.902

3.3　在学者と在職者の成績分析

在学者と在職者の受験者数比率はそれぞれ 46%、54％で、前者 136 名、後者 161 名である。また、在職者のうち、公務員の人数が最も多く、総数 149 名で、そして民営企業 8 名、高校教師 4 名となっている。

全体として、在学者の成績は在職者より優っている。在職者のう

[20] Biserial correlation（双列相関）、試験項目の識別度を表わす。数値範囲は-1 から+1 までの間とする。問題（正解）の Biserial 値は正数でなければならない。+1 に近いほど識別力が高いことを表わす。普通、Biserial 値が 0.1 より低い試験項目の場合は、識別度が悪く、項目を修正或いは削除しなければならない。数値が 0.3 より高い試験項目ほど識別度が高いことを示す（林清山、1992）。Mean Biserial はテストにおける全体の試験項目の双列相関の平均値である。

[21] Cronback's Alpha は、テストの内部一致性（internal reliability）を検証する数値の一つで、テストの信頼性を分析する上での方法の一つとされる。数値範囲は 0 から 1 までの間とする。1 に近いほどテストの安定性、信頼性及び推測可能性が高いことを示す。一般に、専門のテスト機構で作成された利害関係の高い、標準化されたテストのアルファ値は 0.9 以上がいいとされる。利害関係の低い標準化されたテストのアルファ値は、少なくとも、0.80 または 0.85 がいいとされる。授業科目のテストでは、0.70 以上がいいとされる（Wells & Wollack、2003）。

ち、高校教師の成績が最も良く、三項目テスト平均点の合計は 256 で、民営企業の在職者がそれに次いで、三項目テスト平均点の合計は 214 である。この二グループのそれぞれの平均点の合計は、在学者より 28 点以上高いが、受験者数が少なく、グループとしての代表性が貧弱になるため、過大解釈してはならないと考える。

三項目テストのうち、在学者も在職者も、聴力の成績が最も良く、語彙・読解が続き、用法が最後に並んだ。この成績順は全受験者の三項目テストの場合と同じである。特に在学者は聴力が優れており、用法、語彙・読解より平均 10 点高い。在職者は聴力が他の二項目テストより平均 4 点高い。

用法と語彙・読解の二項目テストを比べると、公務員は、語彙・読解の成績が良く、用法より平均 2 点高い。教員は、用法の成績が良く、語彙・読解より平均 2 点高い。民営企業の在職者は、この二項目テストの差異は平均 0.5 点未満で、ほぼ肩を並べている。詳しくは表 3.4 を参照されたい。

表 3.4　在学者と在職者の平均点比較表

在学/在職	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
在学者	136	46%	68.81	58.46	58.66	185.93
在職者	161	54%	61.64	56.14	57.98	175.76
公営	149	50%	59.93	54.80	56.83	171.56
民營	8	3%	79.25	67.25	67.50	214.00
教師*	4	1%	90.00	84.00	82.00	256.00

* 高校教師

3.4　教育バックグランドによる分析

3.4.1　専攻分野による分析

全受験者を専攻分野によって分析すると、「非日本語専攻」グループの人数が多く、日本語専攻と非日本語専攻の比率は 4 : 6 となっている。三項目テスト平均の合計から見ると、「日本語専攻」グル

ープの成績が最も良く、合計 216 点で、非日本語専攻グループより 60 点近く高い。

「非日本語専攻」グループの中で、人数の比率から見れば、「その他の外国語」、「理数化学情報工程」、「管理、財務金融関連」及び「法政関連」という教育バックグランドの受験者が最も多い。三項目テストの平均点については、「農林漁牧関連」が最も良く、合計が 195 点となっている。「管理、財務金融関連」、「社会、経済、心理関連」がその後に続いており、合計はそれぞれ 184 点と 176 点となっている。詳しくは表 3. 5 を参照されたい。

表 3. 5　日本語専攻及び非日本語専攻の平均点比較表

専攻学科	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
日本語専攻	119	40%	78.23	69.16	68.72	216.11
非日本語専攻	178	60%	56.03	49.21	51.32	156.56
その他の外国語	38	13%	51.61	42.16	42.84	136.61
文、史、哲、教育	13	4%	60.00	54.38	53.38	167.77
マスメディア関連	3	1%	63.00	40.67	42.67	146.33
社、經、心理関連	9	3%	66.00	57.33	53.11	176.44
法、政関連	26	9%	51.35	47.35	50.08	148.77
管理、財務金融関連	32	11%	64.44	58.03	61.50	183.97
理數化学情報工程	33	11%	51.39	44.21	49.15	144.76
醫藥衛生生命科學	10	3%	48.60	52.80	51.20	152.60
農林漁牧関連	6	2%	65.67	63.83	65.33	194.83
無	1	0%	26.00	37.00	34.00	97.00
その他	7	2%	59.29	48.86	51.14	159.29

3.4.2　学年別による分析（在学者）

在学者学年別で分析すると、3 年生と 4 年生の参加比率が最も高く、それぞれ 3 割と 5 割となっている。三項目テストの成績は修業した年数に正比例して、4 年生グループの成績が最も良く、三項目テストにわたり、すべて 66 点を超えている。各言語技能別で大学生

の伸び具合を見た場合、「3 年生に昇級」する段階での向上幅が大きく、各項目の平均の差異はそれぞれ 11 点以上の開きが見えるのに対し、「4 年生に昇級」する段階での向上幅は縮まり、各項目の平均の差異は 7~8 点に止まっている。

項目テスト別の平均点を高低順に並べて見た場合、各学年とも、聴力が最も良く、用法と語彙・読解の平均より 6~10 点高い。特に 3 年生の三項目の高低順は全受験者の場合と同じく、聴力が最も高く、語彙・読解が続き、用法が最後に並んだ。その他の二グループも聴力が最も良いが、用法が後に続き、語彙・読解が最後という順序になっている。

各学年のうち、3 年生の各項目の平均点はすべて 58 点を超え、全受験者のパフォーマンスに最も近接している。詳しくは、表 3.6 を参照されたい。

表 3. 6　在学者各学年の三項目平均点比較表

学年	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
2 年生	23	18%	53.17	46.83	46.35	146.35
3 年生	45	34%	68.47	57.71	58.84	185.02
4 年生	60	46%	76.48	66.05	65.87	208.40

有効サンプル数：128 人

3.4.3　学習時間数による分析（在職者）

日本語の学習時間数によって在職者を集計したところ、「450 時間以上」グループの人数が最も多く、有効サンプルの 40%を占めている。次に多いのは「250~300 時間」グループで、34%を占めている。各グループの平均点については、「350~400 時間」グループを除き、すべてのグループが学習時間数に比例していることが分かった。「450 時間以上」グループの成績が最も良く、三項目とも平均 68 点を超え、合計が 211 点となっている。

各グループのうち、「400~450 時間」グループの三項目はそれぞ

れ平均 59 点を超え、全受験者の成績に最も近接しているとともに、三項目の合計が 184 点となり、教育部の公費留学の合格基準をクリアしている。各グループの成績については、表 3.7 を参照されたい。

表 3.7 学習時間数による各グループの三項目平均点比較表

學習時間數	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
250~300 時間	51	34%	48.63	43.02	44.43	136.08
300~350 時間	18	12%	61.67	54.44	56.56	172.67
350~400 時間	12	8%	51.58	51.75	54.17	157.50
400~450 時間	12	8%	62.83	59.33	62.00	184.17
450 時間以上	59	39%	74.12	67.53	69.56	211.20

有効サンプル数：152 人

3.4.4　年齢層による分析

年齢から見た場合、受験者の年齢分布が、18 歳から 60 歳まで広がっている。そのうち、「20~24 歳」の受験者数が最も多く、44%で、ほとんどが在学生である。在職者がほとんどである 25 歳以上の各グループの中で、「25~29 歳」と「30~34 歳」グループの受験者数比率がそれぞれ 14%と 11%で、35 歳以上の各年齢層の受験者数比率は、年齢増になるにつれ緩やかに減りつつあるが、日本語学習者の年齢層が若年から高年まで続いていることを示している。図 3.8 を参照されたい。

各年齢の三項目テストの成績について、「60 歳以上」グループが最も良く、その後に「20~24 歳」、「25~29 歳」、「35~39 歳」と順を追って並んでいる。グループ別の三項目平均の合計点は、それぞれ 204、189、182、と 180 である。全体にわたり、人数が比較的少ない「15~19 歳」と「60 歳以上」の二グループを除くと、各年齢層の成績高低順がほぼ年齢の高低に反比例しているという現象が見られる。これは外国語学習において、年齢的な要素が大きいことを表わ

している。

表 3.8 各年齢層の三項目平均点比較表

年齢層	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
15~19 歳	5	2%	54.60	44.60	44.80	144.00
20~24 歳	130	44%	69.83	59.23	59.52	188.58
25~29 歳	42	14%	64.90	57.83	59.62	182.36
30~34 歳	33	11%	64.76	54.67	56.85	176.27
35~39 歳	20	7%	62.75	57.10	60.50	180.35
40~44 歳	27	9%	59.22	56.70	57.78	173.70
45~49 歳	13	4%	55.15	52.85	53.54	161.54
50~54 歳	17	6%	57.41	54.94	57.29	169.65
55~59 歳	6	2%	41.67	45.17	42.67	129.50
60 歳以上	4	1%	67.75	67.00	69.00	203.75

項目テスト別で見た場合、聴力につき、「20~24 歳」グループの成績が最も良いが、用法、語彙・読解の二項目については、「60 歳以上」の成績が最も良い。各年齢層において、用法と語彙・読解の成績差が小さく、わずか 3 点以下に止まっている。そして、聴力だけを見た場合、大まかに「35~39 歳」を境に、それより若い各年齢層では聴力が他の二項目と大きく差をつけているが、それより高い年齢層では聴力と他の二項目との成績差が 3 点以下に縮まった。

上述のことから、中年以前は聴解能力が用法と語彙・読解より目立っているが、中年以降は三種の技能がバランスよく伸びていると言うことができる。ただし、一部の高年齢層の人数が少ないため、この現象が安定しているかどうかは、さらなる研究を待たねばならない。各年齢層の平均点については表 3.8 を参照されたい。

3.4.5　職場での使用頻度による分析

回収したアンケートによって、在職者の聴く、話す、読む、書く

及び用法の五つの言語技能の職場での使用頻度を集計した結果、「読む」技能が最もよく使われ、「頻繁」と「時々」が合わせて 60%弱となるのに対し、使用頻度が最も低いのは「書く」技能で、約 30%に止まっている。「聞く」、「用法」と「話す」の三つの技能がそれぞれ「頻繁」と「時々」を合計した使用比率は約 45%前後となっている。全体として、職場で上述の五つの言語技能がすべて使用されてはいるが、「頻繁」使用の比率はそれほど高くないことが見られる。さらに、五つの言語技能の使用頻度は、筆記テストの成績に正比例していることが見られる。即ち、五つの言語技能にわたり、「頻繁」使用者の三項目成績の合計は、すべて 200 点を超え、「時々」、「たまに」の使用者より成績が良い。その中で特に、発表技能[22]使用者の成績が最も良く、「書く」グループの三項目平均の合計が 235 点、「話す」グループが 231 点で、それに続いて、「聴く」の 222 点、「用法」の 212 点、「読む」の 201 点となっている。詳しくは表 3.9～3.13 を参照されたい。

表 3.9　在職者**聞く**技能の使用頻度と平均点の比較

使用頻率	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
頻繁	14	9%	81.43	67.86	73.00	222.29
時々	52	35%	61.79	54.81	56.00	172.60
たまに	84	56%	57.63	54.68	56.64	168.95

有効サンプル数：150 人

表 3.10　在職者**用法**技能の使用頻度と平均点の比較

使用頻率	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計

[22] productive skill（発表技能）、話す、書くのような技能を言う。聴く、読むのような receptive skill (受容技能)に相対している。

頻繁	14	9%	79.36	65.57	66.57	211.50
時々	56	37%	61.04	56.16	58.46	175.66
たまに	80	53%	58.31	54.13	56.08	168.51

有効サンプル数：150 人

表 3.11　在職者**読む**技能の使用頻度と平均点の比較

使用頻率	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
頻繁	31	21%	71.81	63.35	65.61	200.77
時々	54	36%	59.81	54.48	57.15	171.44
たまに	65	43%	57.51	53.65	54.95	166.11

有効サンプル数：150 人

表 3.12　在職者**話す**技能の使用頻度と平均点の比較

使用頻率	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
頻繁	16	11%	84.56	72.25	74.50	231.31
時々	48	32%	63.25	55.29	58.25	176.79
たまに	86	57%	55.87	53.29	54.70	163.86

有効サンプル数：150 人

表 3.13　在職者**書く**技能の使用頻度と平均点の比較

使用頻率	人数	人数%	聴力平均	用法平均	語彙・読解平均	三項目合計
頻繁	10	7%	87.10	73.20	75.00	235.30
時々	31	21%	56.45	51.48	54.00	161.94
たまに	109	73%	60.30	55.64	57.50	173.45

有効サンプル数：150 人

4.　結び

二期研究計画において、2005 年 4、5 月に多段サンプリング調査で、全国 37 の大学の日本語専攻学生と非日本語専攻学生、及び 73 の公、民営機構の在職者、合わせて 297 名の受験者を抽出し、テスト実施、アンケート調査を行ない、集計分析した結果が、

下記の通りまとめられた。

なお、紙幅に限りがあるため、学歴別、地域別、性別による分析の報告を省略せざるを得ないことをご了承されたい。

1) 台湾における 18 歳以上の日本語成人学習者（真性初心者を含まず）の日本語能力において、聴力が最も良く、語彙・読解が次、用法が最後に並んだ。平均得点は、それぞれ 65 点、58 点と 57 点となっている。
2) 三項目テストにわたり、平均有意確率（Mean P）が 0.6 前後、平均双列相関（Mean Biserial）が 0.4 以上、アルファ値が 0.9 以上であることは、各項目テストの難易度が適切で、識別度と信頼性が高く、受験者能力を有効に区分することができ、テスト結果は安定性と推測可能性を持っていることを示している。このことから、FLPT 日本語能力テストは日本語能力の評価ツールとして適切であることを表わしている。
3) 三項目のパフォーマンスにおいて、在学者は在職者より優れている。しかし在職者の教師と民営企業の二グループの能力は、在学者を追い越している。項目テスト別の能力において、在学者、在職者共に聴力が最も良く、特に在学者では聴力が他の二項目より平均 10 点以上高く、在職者では他の二項目テストより平均 4 点以上高い。
4) 専攻分野別で見た場合、非日本語専攻グループの人数が多く、全受験者の 60%を占めている。三項目テストの成績では、日本語専攻グループのほうが優れ、三項目平均点の合計が 216 点で、非日本語専攻グループより 60 点高い。非日本語専攻グループの中で、「農林漁牧」、「管理、財務金融」背景者の成績が良く、三項目平均点の合計がそれぞれ 195、184 点となっている。
5) 学年別で見た場合、在学者の成績は、ほぼ修業年数に正比例している。各学年のうち、大学 4 年生の成績が最も良く、三項目それぞれの平均は、すべて 66 点を超えている。言語能力の

伸び具合については、「3 年生に昇級」の段階の向上幅が最も大きく、「4 年生に昇級」の段階がそれに続いている。テスト項目別の平均点につき、各学年はすべて聴力が最も良く、他の二項目より 6~10 点高い。各学年のうち、大学 3 年生の成績は三項目とも平均 58 点を超え、受験者全体のパフォーマンスに最も近接している。

6) 在職者の学習時間によって分析した結果、その平均点が学習時間に正比例していることが分かった。「450 時間以上」グループは各項目とも 68 点を超え、成績が最も良い。「400~450 時間」グループの各項目テストはすべて平均 59 点を超え、受験者全体の各項目の平均に最も近接しているとともに、合計が 184 点で、教育部の公費留学の合格基準をクリアしている。

7) 年齢層によって分析したところ、「60 歳以上」グループの成績が最も良く、その後に「20~24 歳」、「25~29 歳」、「35~39 歳」と順次に続いている。三項目の合計は 4 グループとも 180 点を超えている。年齢層別の成績は概ね年齢の高低に反比例している現象が見られ、言語学習において年齢的な要素が大きいことを物語っている。項目テスト別で見た場合、聴力では「20~24 歳」グループの成績が最も良く、用法、語彙・読解の二項目では、「60 歳以上」の成績が最も優れている。さらに聴力だけを見た場合、大まかに「35~39 歳」を境に、それより若い各年齢層では聴力が他の二項目と大きく差をつけているが、それより高い年齢層では聴力と他の二項目との成績差が 3 点以下に縮まった。このことから、中年以前は聴解能力が用法と語彙・読解より目立っているが、中年以降は三種の技能がバランスよく伸びていると言える。これは、学習方法の多様化に関係していることに原因があるように思われる。特に 1993 年から、日本のテレビ番組・映画と音楽に対する政府の解禁政策[23]、及びインター

[23] 「日台文化芸能の出来事」（台湾資料センター）を参照。

ネットの著しい発展により、日本語学習者に、従来の教室以外の学習環境、資源が提供されていることが、学習者の日本語能力の向上に寄与していると考えられる。

8) 日本語の聴く、話す、読む、書く及び用法の五つの技能が、すべて職場で使われているが、それぞれ「頻繁」に使用している比率は 20%以下に止まっている。そのうち、「読む」技能の使用頻度が最も高く、「書く」技能の使用頻度が最も低いことが見られる。テスト結果の成績が使用頻度に正比例していることが分かった。即ち、各技能にわたり、使用が頻繁である各グループの三項目平均の合計は、すべて 200 点を超え、「時々」と「たまに」のグループより 29 点以上の成績差が見られる。使用が「頻繁」である各技能の中で、発表技能の使用者の合計点が最も高いことが見られ、「書く」と「話す」の二グループのそれぞれの合計は、230 点を超えている。

【参考文献】

J・D ブラウン（2003）『言語テストの基礎知識』大修館書店

格式化

ライル　F　バックマン　（1997）『言語テスト法の基礎』東京　みくに出版

余民寧（2002）　『教育測驗與評量－成就測驗與教學評量』（第二版）心理出版社

王文科（1999）　『教育研究法』　五南圖書出版公司

林清山（1999）　『心理與教育統計學』東華書局

郭生玉　（1997）　『心理與教育測驗』　11 版　精華書局

語言訓練測驗中心（2004）＜台灣地區大學校院歐日語主修生語言能力評量研究計畫－日語評量報告＞　語言訓練測驗中心

語言訓練測驗中心（2004）＜英語能力測驗成績統計報告－民國 91~92 年＞　語言訓練測驗中心

語言訓練測驗中心（2002）＜英語能力測驗成績統計報告－民國 89~90 年＞語言訓練測驗中心

語言訓練測驗中心（2001）＜全民英語能力分級檢定測驗－90 年中級測驗 初、複試成績統計報告＞語言訓練測驗中心

語言訓練測驗中心（2001）＜全民英語能力分級檢定測驗－90 年第一次中高級測驗 初、複試成績統計報告＞ 語言訓練測驗中心

語言訓練測驗中心（2000）＜全民英語能力分級檢定測驗－89 年第一次中級測驗 初、複試成績統計報告＞ 語言訓練測驗中心

語言訓練測驗中心（2000）＜全民英語能力分級檢定測驗－89 年第二次中級測驗 初、複試成績統計報告＞ 語言訓練測驗中心

語言訓練測驗中心（1999）＜英語能力測驗成績統計報告－民國 84~87 年＞ 語言訓練測驗中心

Borg, W.R., & Gall, M.D. （1989）*Educational Research: An Introduction (5th ed.)* New York: Longman.

インターネット上の資料

台湾資料センター（2005）＜日台文化芸能の出来事＞ http://www.roc-taiwan.or.jp/data/200506tj.pdf

教育部公告 （2000）＜八十九年公費留學考試語言鑑定機構及語言能力合格標準＞ http://www.edu.tw/EDU_WEB/EDU_MGT/BICER/EDU7954001/aa/aa-15.htm

教育部公告 （2000）＜八十九年公費留學考試留學國語文證明＞ http://www.edu.tw/EDU_WEB/EDU_MGT/BICER/EDU9566001/c23/c230303-3.htm**Wells, C.S., & Wollack, J.A.**（2003） *Guide to Understanding Test Reliability.* http://wiscinfo.doit.wisc.edu/exams/Reliability.pdf