

外国語習得における筆記試験の役割

浅野 真紀子

サンフランシスコ州立大学

masano@sfsu.edu

The Role of Written Examinations in Foreign Language Learning

Makiko Asano

San Francisco State University

Abstract:

In this paper I argue that written examinations have a positive influence on learning foreign languages. While language is a means of communication, language learners are often encouraged not to communicate using language per se, but rather to perform problem-solving tasks through written drills and exercises. When the purpose is to enhance communication, are written examinations necessary in teaching a language? Do they have a positive influence in learning a language? In order to investigate the usefulness of written examinations, an experiment was conducted involving 41 university students in beginners' Japanese language classes. The subjects were divided into two groups: the experiment group was given three written quizzes after the *hiragana* stroke order was introduced. The control group was not given such quizzes. These two groups were examined for three months, their rate of retention of *hiragana* stroke order statistically monitored and analysed. The data revealed that the experiment group retained the *hiragana* stroke order better than the control group.

1. はじめに

語学教育に携わるものであれば、筆記試験・小テスト、国によってはクイズなどと呼ばれる評価手段が、学習者にどういった影響を与えるか、学習効果にどういった役割を果たすのか、果ては本当に必要なのかなどと考えたことがあるだろう。大学などの教育機関での語学教育を考える時、学習者の言語習得進度を測るため「話す、聴く、読む、書く」の四技能を網羅した試験が行われるのは珍しくない。さらに、時と場合に応じて単語テスト、漢字テスト、活用テストといった語学習得の一側面に焦点を当てた小テストなども併用されるのが現状ではないだろうか。しかし、人と人と

のコミュニケーションを図ることを目的とする言語の教育に筆記試験は本当に必要なのだろうか。

言語は本来意志伝達のためのツールである。外国語を学ぶ学習者が机に向かって黙々と試験問題を解いている図は、言語習得の本来の目的とは何かはずれていることは否めないだろう。また、筆記試験を受ける学習者たちの心理的圧迫感に関する問題や、試験そのものの信頼性・妥当性の問題も無視できない (Backman 1990, Cohen 1980, Davies 1990)。その一方、大学などの教育機関における外国語教育の現状を考える時、個々の学習者はもとより、教室別、レベル別、果ては学校別や、地域別といったさらに大きな視点から語学指導を見据えるための評価手段として、筆記試験・小テストの効用には計りしれないものがある。

それでは、筆記試験はこうした評価ツールとして以外にどんな効用があるのだろうか。言い換えると、筆記試験が学習者の言語習得の本質に正の効果をもたらすか否かを具体的な例を用いて考えるのが本稿の目的である。単に成績を出すための手段として以外に、筆記試験が言語習得そのものに正の効果をもたらすという結果が出れば、いささか場違いとも思われるような筆記試験を用いることに大きな意味を与えることができるからである。

2. 筆記試験の是非

ここではまず、筆記試験を行うことの是非についてもう少し深く考えてみたい。筆記試験の利点でまず思い付くのは、上に述べたように、評価手段としての効用である。筆記試験の多くは結果がはっきりと点数に換算できる。その善し悪しは別にしても、数字は簡単に学習者の能力と関連させて考えることができ、教師にとっても学習者にとっても学習進度が理解しやすい (Cohen 1980)。また、教師にとって学期末の総合評価を出す際にも、数字に換算した試験などの成績は大変便利なものである。学習者の立場から考えてみても、例えば米国においての「85点なら成績はB」など

といった概念は受け入れやすく、自分のパフォーマンスに対する教師の評価を理解するのに混乱が生じにくいだろう。さらに、学習者自身も筆記試験などがなければ、どうやってこの先生は自分たちの成績をつけるのだろうと逆に不安になるのではないだろうか。

また、出題範囲やその内容などを予告した上で筆記試験を行う際、学習者の意識が高まるという利点もあるだろう。つまり「これは試験に出るのだから、大事なのだ。」といったような、ある学習項目に対する関心、注目度が上がるということである。さらに、試験や小テストは学習者たちが学んだことを正しく理解、習得しているか確かめるための手段にもなる (Cohen 1980)。試験問題に取り組み、その後間違いを確認し、正しい答えを学ぶことで、復習のためのツールとしても使えるのである。

また、筆記試験や小テストが学習者の勉強意欲を促進するという点も忘れてはならない。勉強意欲促進というと少し言い過ぎかもしれないが、「明日試験があるから勉強する」、逆に「来週は試験がないから週末は遊ぶ」といった公式は多くの学習者に当てはまるのではないだろうか。試験がなければ遊びに行っていたであろう学習者が、翌日の試験のために机に向かうのは決して見慣れない光景ではない。

もちろん、筆記試験にはこうした利点がある反面、欠点と思われる面も多々見つかる。まず、筆記試験・小テストを実施するという事は、例えば会話の練習時間や講義時間といった語学教育に欠かせない大切な要素と引き換えに行われる。その上、残念ながら多くの学習者は試験などは歓迎しないのが現状である。試験の時に不必要なストレスを感じ、思うように実力が発揮できない学習者も少なくない (Cohen 1980)。それに関連して、試験そのものの評価基準や結果に対する信頼度の問題もある (Backman 1990, Davies 1990, Cumming & Berwick 1996)。例えば、教室活動では上手に意志疎通ができるのに試験をしてみるとあまりできない学習者がいると、果たしてクラスパフォーマンスと試験の評価のどちらがその学習者の本来

の能力なのか迷ってしまう。

また、教師にとって、試験作りは大変気を遣う作業でもある。教えたことを網羅しながら、評価基準、出題内容など信頼するに足る公平な試験を作るのには時間も手間もかかる。採点する時に至っても同じである。正誤の確認、解答内容の把握はもとより、個々の間違いから学習させるためにコメントを書き入れるなど、やはり時間も手間もかかる。

さらに、上に述べたように、試験や小テストなどを実施することは本来の語学習得の目的とは離れてしまうという懸念もある。特に大学レベルでの日本語教育は、ともすれば他の歴史や経済学などといった非語学クラスと同様に扱われるため、試験という手段がカリキュラムや採点の格の一つとしてシラバスが作られることが多い。これが本末顛倒となって、コミュニケーションツールとしての言語を学ぶという本質的な意識を忘れ、いい成績をとりたいたから試験のために勉強するということになってしまう可能性すらある。

さて、こうした欠点が多々ありながらも、やはり筆記試験や小テストを廃止するという決断はなかなかできないのが普通だろう。学校教育においては、何らかの評価を個々の学習者に与えることが常識かつ必須だからだ（逆に、家庭教師など、成績を出すことを必ずしも必要としない場面では、筆記試験を行う率は非常に低いのではないだろうか）。仮に、試験や小テストなどの主な効用が単に成績を出すことだとしたら、果たして多くの犠牲を払ってまで行うことが本当に必要なのか、原点に戻って一度考えてみるべきではないだろうか。反対に、もし筆記試験などが学習者の言語習得促進に良い効果をもたらすのであれば、これを行う意義は大きいと言える。

3. 実験

この研究では、ひらがなの筆順習得という学習項目を例に用い、筆記試験の有無により学習内容の定着度が変化するか否かを調べた。無作為の二

グループに分かれた被験者は、はじめにひらがなの筆順に関するアンケートを受ける。これは、後に続く実験の基準、もしくはスタートラインともなるもので、13のひらがなの筆順を直感で答えるものである。この後、両グループは同じスケジュール、同じやり方でひらがなを紹介されるが、この際一つのグループの被験者のみが、事前に予告を受けた上で（つまり教室外で学習する機会を与えられた上で）筆順に関する小テストを三回受ける。二つのグループの違いはこの小テストの有無だけである。その後、筆順の定着度を調べるため、導入が終わって一ヵ月後と三ヵ月後に両グループに予告なしで筆順小テストを行った。

上述のように、この実験は二グループのうちの一グループにだけ筆記試験を課すという倫理面でデリケートな問題をはらむため、実験の題材として、学習者の日本語習得の基盤がかかってしまうというようなものは避けた。そこで仮に二つの被験者グループの学習項目の定着度に違いが出たとしても、後々の日本語学習に支障が少なく、進度の遅れも修復がたやすいと思われる「ひらがなの筆順の習得」という学習項目を選んだ。

ここで、具体的な実験内容に移る。被験者は、アメリカの大学で初級日本語の一学期目を履修する41名の学習者である。¹ この41名は17名からなるグループ（以後グループAと呼ぶ）と24名からなるグループ（以後グループBと呼ぶ）の二つに分かれる。この二つのグループの人数の違いは、同じ学期に開講した日本語の二つのクラスを被験者グループとして選んだため、それらのクラスの履修人数の差がでたものである。この二グループは一回の授業が50分、週5回の授業を15週間にわたって受講した。

このA・Bの両グループは通常シラバスの説明など、オリエンテーションが行われる日である学期の初日、つまり被験者たちが日本語に触れたこともない時点であるが、この日にひらがなの筆順に関する直感力を調べるアンケートを受けた。（1）はそのアンケートからの抜粋であるが、これはその後ひらがな筆順の導入後に行う二回の追跡小テストの結果と比較し、

定着率を測るための基準のテストともなる。そこで、この初日に行ったアンケートを以下では基準テストと呼ぶ。

(1) 基準テスト抜粋

Native language: _____ Name: _____

Have you ever studied how to write Hiragana?: Circle one Yes No

Please don't panic! This is not a quiz or exam. We would just like to know how you would intuitively write Japanese characters without being taught stroke order.

Example

Completed	1 st stroke	2 nd stroke
こ	こ	こ

[Instruction] Please use a pen in any other color than black, if possible. Complete characters are shown in the leftmost column and, depending on the number of strokes, a certain number of characters in shaded lines are given in right columns. Trace the line of the indicated stroke in each box, showing how you'd write the given character. Don't think. Follow your instinct.

Completed	1 st stroke	2 nd stroke	3 rd stroke
あ	あ	あ	あ

Completed	1 st stroke	2 nd stroke	3 rd stroke	4 th stroke	5 th stroke
が	が	が	が	が	が

この基準テストでは、被験者は氏名、母語、そして日本語を過去に勉強したことがあるかどうかの質問に答える。氏名を聞いたのは、この後の二回の追跡調査テストとの差を個々の被験者別に調べるためである。母語を聞いたのは、漢字圏の被験者と非漢字圏の被験者とを比べた時、例えば漢字圏から来た被験者たちがそうでない被験者グループより点がよいなど、この基準テストの結果に統計的な違いが見られれば、これらのグループを一括に扱えない可能性があるからである。同じ理由で、過去に日本語を学習したことがあるかどうかも聞いた。例えば、過去にひらがなの書き方を習ったことのある被験者たちが、それまでひらがなを見たこともない被験者

たちより基準テストでいい結果が出るなど統計的な差が見られれば、この二つの被験者グループの学習進度を同じように測ることはできないというわけである。

基準テストの質問では(1)にあげたひらがなを含め、全部で13のひらがな「あ、お、が、き、せ、な、ぬ、ふ、ほ、む、も、や、よ」の筆順を聞いている。これらを選んだ理由は、多くのひらがなを出題して基準テストを煩雑にしないようにするためと、そうするには、同じような筆順のものを重複して聞かず、違うタイプの筆順のひらがなを出題しなければならないということである。例えば、多くの被験者は「た」の最初の二画を「な」の最初の二画と同じ筆順で書くであろうと想定し、「た」に比べて三画目、四画目の筆順に被験者によって違いが出そうな「な」の方を採用した。こうした配慮の結果、例外的、または特異な筆順のひらがなを含めて出題数が13となった。例えば、交差する縦線と横線があれば通常横線から書くのに対し、第一画を縦線から始める例外的な筆順の「も」や、通常左から右へ仮名を書き進めるのに対し、二画目、三画目を右から左の順に書く「や」などが問題に含まれている。²

(1)の[Instruction]にある通り、基準テストで被験者に要求されるのは、一番左の欄にある完成したひらがなの文字を見、その右隣にある網かけ文字で示されたひらがなの一つ一つに自分が書くであろう一画一画を直感的に書き足していくことである。基準テストには例として「こ」を使った解答の仕方をあげてある。

被験者はこの作業を上にあげた13のひらがなに対して行うわけだが、特に時間の制限を設ける理由もなかったもので、全部できた時点で提出するという形をとった。³以上が、ここで基準テストと呼ぶ実験の第一ステップである。この後、被験者たちは授業の二日目から五日間にわたってひらがなの音と文字、そしてその筆順を学んでいく。

実際の教室指導は、ひらがなを含めた日本語の文字の起源の説明に始ま

り、「外来語はカタカナで書く」など、各々の文字の使用対象を述べた後、ひらがなを書く際の一般的な規則を導入した。(2)に、教室で指導したひらがなを書く際の一般的規則をあげる。

(2) ひらがなの筆順に関する一般規則

- a. 一つの仮名は左から右へ書く。例：「い」
- b. 一つの仮名は上から下へ書く。例：「こ」
- c. 縦線と横線が交差する場合は横線を先に書く。例：「せ」
- d. 濁点などの点は最後に書く。例：「お、が」

この後、一つ一つのひらがなの筆順と読み方を紹介していったが、具体的なひらがなの筆順導入のスケジュールは(3)の通りである。

(3) 教室指導スケジュール

- 初日： 基準アンケート実施
- 二日目：日本の文字の種類・起源、使用対象、筆順の一般規則、ひらがな「あ～そ」
- 三日目：ひらがな「た～ほ」
- 四日目：ひらがな「ま～ん」
- 五日目：濁音・半濁音
- 六日目：長音、拗音、撥音、促音の発音と書き方

二日目から四日目に行った個々のひらがなの導入では、一回の授業で紹介する約 15 個に計二十分から二十五分を使い、一つ一つのひらがなの筆順と音を紹介。また、実際に被験者が紹介されたひらがなを書く練習も行った。中でも「も」や「や」などの一般規則に反するようなひらがなの筆順に関しては少し長めの時間を使って丁寧に導入した。さらに、五日目には濁音、半濁音の発音と書き方を紹介し、六日目には直接筆順とは関わりがないが、長音、拗音、撥音、促音の発音と書き方の注意事項を導入した。

こうした導入・練習に伴い、学習者は毎日ひらがなを書く宿題を与えられた。この宿題をする際、筆順を示してあるワークブックを使うため、学習者は教室外でも個々のひらがなをどう書くかを目にするようになる。さ

らに学習者は教科書に紹介されている筆順のページはもちろん、インターネット上のひらがなを紹介するサイトで、個々のひらがなの書き方を動画で見て復習できるような情報も与えられている。

こうした教室内での導入、教室外での宿題などはグループA、グループB共に、極力同じ長さの時間を使い、同じ説明方法でなされた。従ってひらがなの筆順に関する指導、学習者に与えられた課題はA・Bの両グループにとって、限り無く近いものであったと言える。

さて、筆順の導入・学習が終わった後、両グループに与えられる課題の一つだけ違いが出てくる。それが三回の小テストである。グループAだけが、事前に予告を受けた上で、つまり教室外で勉強するチャンスを与えられた上で、筆順の小テストを三回受けた。⁴この小テストはひらがなを全て導入し終わった翌日から、一回目は「あ」から「そ」まで、二回目は「あ」から「ほ」、三回目は「あ」から「ん」までと濁音という出題範囲で行った。⁵形式は、上記(1)に挙げた基準テストと同じで、網かけ文字で書かれたひらがなの一画一画を筆順通りになぞるというものだが、問われたひらがなは出題範囲により、基準テストとは異なっている。

一回の小テストで聞いたひらがなの数は七個であったが、この小テスト実施において重要なことは、その内容ではなく、グループAの学習者は教室外でも勉強する動機を与えられたということである。グループAがこうした三回の小テストを受ける一方、グループBにはこの小テストが課せられなかった。つまり、ひらがなの宿題を提出した後、グループBは特に教室外で、限られた時間までに筆順を勉強しようという強い動機がなかったかもしれない。二つのグループの違いはこの三回の小テストの有無だけである。

さて、ここからが実験の本題になるが、筆順導入時に三回の小テストを行ったグループAと行わなかったグループBに筆順の定着率の違いが出るかどうかを見るため、筆順導入時より一ヵ月たった時、両グループに予告

なしで、ひらがなの筆順の小テストを与えた。さらに長期に渡っての定着率を見るため、ひらがな導入時より三ヵ月目にも両グループに予告なしで同じ小テストを行った。一ヵ月目のテストも三ヵ月目のテストも、内容は初日に行った基準テスト（（1）参照）と同じ 13 のひらがなの出題で、答え方の形式も同じである。

4. 両グループの性質

以上が、実験の詳細であるが、これらの結果の分析に移る前に、この実験以外での両グループの教室活動を比較した結果を述べておきたい。もともと同じレベルのクラスであるが、この二つのグループに通常から優劣の差があれば、筆順テストに関する実験結果の比較もあまり意味をなさないからである。

まず、各課を学習し終わった後に行う試験の結果を両グループで比べた。これは「読む、書く、聞く」の技能の学習進度を測る総合的な試験で、被験者たちが受講した一学期を通じて六回行われた。六回の試験結果を合わせて 100 点満点に換算した結果が（4）である。

（4）課毎の試験の結果

	平均点	標準偏差値
グループ A	85.98	12.79
グループ B	86.64	10.48

（4）の結果を t 検定を使って比較したところ、t 値が 0.175、有意水準 5% で、二つのグループの試験結果に統計的な差は見られないという結果が出た。つまり両グループとも各課毎の試験の結果には優劣の差はないということである。

また、一学期を通じて二十八回行った小テスト（筆順の小テストとの混同をさけるため、以下クイズと呼ぶ）の結果も比べた。これは語彙や漢字

のクイズ、動詞や形容詞の活用クイズなどで、一回につき約三分から七分かけて行ったものである。(5)にそれらを総合して10点満点に換算した数値をあげる。

(5) クイズの結果

	平均点	標準偏差値
グループA	8.24	1.76
グループB	8.04	1.26

(5)の結果を基にしてt検定を行ったところ、t値は0.394であった。クイズの成績の結果も有意水準5%で、二つのグループに統計的な違いは見られないことがわかった。

さらに、二時間半に渡って行われた学期末の試験結果でも二つのグループを比べた((6)参照)。

(6) 学期末試験の結果

	平均点	標準偏差値
グループA	77.34	1.88
グループB	76.78	1.48

(6)の結果を統計的に見ると、t値が0.102で、期末試験の成績も有意水準5%で、二つのグループの成績に違いがないと言えることがわかった。つまり、課毎の試験、クイズ、期末試験では両グループの成績に統計的な優劣の差は見られず、A・Bの二グループは日本語習得、その定着率に関して本質的に同じパフォーマンスをしている被験者集団であると言える。

5. 実験結果

上記を踏まえた上で、実験結果の報告に移る。はじめに、漢字圏、非漢字圏といった被験者の母語背景の違いが初日に行った基準テストに影響が

あったかどうかについて述べる。これは 11 名の漢字圏の学習者の方が 41 名の非漢字圏の学習者よりひらがなの筆順に対する直感がある、と言えることが統計的にわかった。^{6,7} 普段から漢字を使っているという言語背景があると、ひらがなの筆順も習う前からある程度推測できると言ってよいだろう。従って、上に説明してきた A・B 両グループの被験者の中には漢字圏から来た学習者 11 名は含まれていない。

次に、日本語学習の経験の有無が基準テストの結果に違いをもたらしたかどうかについてであるが、こちらは漢字圏から来た被験者を除く 41 名のうち、14 名の日本語学習経験のある被験者も 27 名のない被験者も統計的に違いがないことがわかった。^{8,9} 過去にいくらか日本語を学習したことがある被験者も、筆順に関しては日本語未経験の被験者と同じレベルの知識、または直感しかもっていない、といういささか残念な結果ではあるが、過去の日本語学習経験の有無が基準テストの結果に統計的な差をもたらさないという結論に従って、A・B 両グループの被験者には過去に日本語を学習した経験をもつ者も含まれている。

それでは、ひらがな筆順導入より一ヵ月たった時に両グループに予告なしで行った小テストの結果に移る。(7) はその結果を 10 点満点に換算したものである。

(7) 筆順導入後一ヵ月の追跡テスト

	平均点	標準偏差値
グループ A	9.37	0.83
グループ B	8.46	1.79

(7) を統計的に見ると、t 値が 2.523、有意水準 5% で、筆順導入時に小テストを受けたグループ A の成績の方がグループ B の成績よりいいという結果がでた。

また、筆順導入後三ヵ月たった時点で行った予告なしの小テストで、さ

らに長期に渡る筆順の定着の具合を調べた結果が（８）である。

（８）筆順導入後三ヵ月の追跡テスト

	平均点	標準偏差値
グループA	9.14	1.12
グループB	8.62	1.38

（８）を統計的に見ると、 t 値は 1.845 となり、三ヵ月後の小テストも有意水準 5 % で、グループ A の成績の方がグループ B の成績よりいいと言えることがわかった。

各課毎に行われる総合試験、学習項目に合わせて行われるクイズ、総合学習進度を見る期末試験、そして筆順導入以前に行った基準テストでは全く統計的な差の見られなかった両グループであるが、筆順に関する小テストを三回行ったか行わなかったかで、その後の定着率の結果に歴然とした差が出た。このデータを見ると、筆記試験は第二節であげたような欠点を数多く持つ反面、学習者の学習事項の定着に確実にいい影響を与えていると言える。

6. 今後に向けて

以上から、筆記試験の実施が学習項目の定着に大きく貢献していることがわかった。今後の課題はまず、この定着率の上昇がなぜ起こるのかを調べて見ることである。学習者が試験の前に勉強することが、長期にわたっての学習項目の定着を促進させるのだろうか。ある項目に関しての試験があるという意識が学習者の注意を喚起するのだろうか。それとも試験を受け、間違いを正すことから定着率が向上するのだろうか。いずれにせよ、筆記試験は成績を出すための道具ではなく、学習効果を上げる道具としても使えるのである。

また、今回の実験は被験者への倫理的配慮により「ひらがなの筆順の習

得」という本来の語学学習の目的である「意志伝達をはかる」という目的からはやや離れたところに焦点を置いて行われた。今後、こうした研究を発展させ、倫理的配慮を欠くことなく、意志伝達を図る手段としての言語の習得を筆記試験が促進するかどうか調査・分析することが望まれる。その結果が本稿と同じく「筆記試験は学習内容の定着に正の効果をもたらす」と出れば、さらにより効果的な筆記試験の方法論について研究を進めたいと思う。

注

- 1 実際には 52 名の被験者がいたが、そのうちの 11 名は母語で漢字を使用し、それにより、ひらがなの筆順に対してもある程度の知識があった。事前調査により漢字を使用しない言葉を母語とする被験者と同等に扱えないことがわかったので、これら 11 名の被験者は実験から除外した。詳しくは第 5 節を参照。
- 2 この結果、全問を日本人が書くような筆順で書くという直感を持った被験者はいなかった。
- 3 全ての被験者が四分以内で基準テストを終えた。
- 4 初日に配るシラバスに、小テストの日程が書いてある他、教室内でも小テスト実施の予告を数度にわたって行った。
- 5 全ての小テストが「あ」から始まっているが、既に過去に小テストの範囲となったひらがな、例えば、「あ～ほ」の範囲の小テストの際の「あ～そ」からの出題は最小限にとどめ、焦点はそれまでに小テストの範囲となっていないひらがな（例では「た～ほ」）である。
- 6 注 1 参照。
- 7 漢字が母語にある被験者 11 名とそうでない被験者 41 名の平均点（10 点満点）と標準偏差値は以下の通りである。

	平均点	標準偏差値
グループ A	6.01	0.96
グループ B	4.16	2.23

t値は 5.18 で、有意水準 5%で、漢字を母語に持つ被験者の方がそうでない被験者より成績がいいという結果が出た。

- 8 過去に日本語を学習した経験者は、最短一ヵ月ほどの大学での日本語教育から最長約二年の高校での日本語教育を受けており、いずれもひらがなの筆順を習った経験がある。
- 9 過去に日本語を習ったことのある被験者 14名とそうでない被験者 27名の平均点、標準偏差値は次の通りである。

	平均点	標準偏差値
グループ A	4.16	2.04
グループ B	5.17	2.24

t値は 1.54 で、有意水準 5%で、日本語学習経験者とそうでない被験者の違いはないということがわかった。

参考文献

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Cohen, A. D. (1980). *Testing language ability in the classroom*. Rowley: Newbury House Publishers.
- Cumming, A., & Berwick, R. (Eds.). (1996). *Validation in language testing*. Clevedon: Multilingual Matters.
- Davies, A. (1990). *Principles of language testing*. Oxford: Basil Blackwell.