

---

# Effects of Time Lag in the Introduction of Characters into the Chinese Language Curriculum

JEROME L. PACKARD  
*Department of Oriental Studies*  
*University of Pennsylvania*  
*Philadelphia, PA 19104*

---

THIS ARTICLE REPORTS THE RESULTS OF A one-year classroom-based study conducted to determine the effects of time lag in introducing Chinese characters into the elementary Chinese language curriculum.<sup>1</sup> Two groups of students, a "lag" (experimental) group and a "no-lag" (control) group, were the objects of this study. The difference between these groups was the amount of time which elapsed before Chinese characters were introduced into their respective Chinese curricula. In the case of the no-lag group, the characters were introduced at the very beginning of the course. For the lag group, the characters were introduced three weeks after the course began. The goal of the study was to determine whether a time lag in the introduction of characters results in short- or long-term differences in the second-language acquisition of Chinese, in areas such as speaking proficiency, listening comprehension, phonetic discrimination, grammar, and in the reading and writing of Chinese characters.

## BACKGROUND

Mastery of written Chinese is usually considered difficult because of the large number of nonphonetic, visually complex symbols which constitute the character orthography of the language. For this reason, many Chinese language programs provide a time lag between the time the course starts and the time the characters are introduced. Such a lag generally lasts from one to four weeks, during which time a phonetic, romanized orthography (*pinyin*, at the

University of Pennsylvania) is used for reading and writing.

Such a delay for the introduction of characters is thought to benefit the acquisition of both literacy and oral-aural skills in Chinese. Walker (p. 77) gives two reasons that the delay may benefit the acquisition of literacy skills: first, it follows the pattern of native acquisition of the writing system, since writing is acquired subsequent to speaking; and second, it allows the student to focus on orthography without having to worry about vocabulary and grammar.

The delay is also thought to benefit the development of oral-aural skills because the students build a foundation in the spoken language by initially concentrating on the sounds, vocabulary, and grammar of the language, without spending time acquiring an unfamiliar, complex orthography. In addition, as Walker notes (p. 77), the delay provides a review of studied material, since the delayed characters represent previously learned oral vocabulary.

Teachers of Chinese lack consensus about the provision of a time lag. Some think that delaying the introduction of characters merely postpones an arduous task and that reading and writing skills may in fact suffer as the result of the time lag for at least two reasons. First, the character curriculum must eventually catch up to the regular curriculum, and so students given a time lag actually have a shorter time period in which to learn the characters. Second, the mutually reinforcing effect of characters and the words they represent may be lost when they are introduced at different times.

## RATIONALE

While a delay in the introduction of characters is believed to benefit the second language acquisition of Chinese, this belief had not yet

been empirically substantiated. It had also not been established that delaying the introduction of characters is not actually *detrimental* to the acquisition of literacy skills. Therefore, I decided to study the effects of introducing the characters at different times to two different groups (in an otherwise identical curriculum), to see if resultant differences would occur in the acquisition of the language.

#### METHOD

*Subjects.* The subjects of this study were students enrolled in a double-credit (two credits rather than one), intensive elementary Chinese course at the University of Pennsylvania during academic year 1988-89. The lag and no-lag groups were two separate sections of the course, which were scheduled at the same time in order to help ensure a homogeneous subject population.<sup>2</sup>

At the beginning of the course, students were asked to fill out a form requesting information on the extent of their prior experience with the Chinese language. The requested information included number of characters recognized, previous schooling in Chinese (formal or informal), speaking and listening ability in Chinese (any dialect), whether Chinese was spoken in the home, and so forth. The students were asked to give a numerical response from 1 to 5 for each question, representing their level of previous exposure to Chinese. The total for each student yielded a numerical measure of previous experience with the language and provided a statistical control for amount of Chinese language background.

In addition, subjects' scores on the verbal portion of the Scholastic Aptitude Test (VSAT) were used as a statistical control on verbal ability. The background and verbal scores for each subject, as well as group means for these variables, are seen in Table I.

*Texts.* During the fall semester, two textbooks were used, *Beginning Chinese* and *Character Text for Beginning Chinese*, both by John DeFrancis (2, 3). In the second (spring) semester, the DeFrancis series was used until about halfway through the semester, when *Read Chinese* by Richard Chang, which uses only Chinese characters, was used.

*Groups.* The course was in session for thirteen weeks each semester. The lag group did not begin to learn characters until the beginning of the fourth week of the course, at which

TABLE I  
Subjects

Subject No.	Background Score	VSAT Score
LAG GROUP		
1	9	560
2	10	660
3	7	520
4	5	730
5	7	600
6	8	660
7	8	660
8	5	620
9	5	550
10	9	720
11	5	700
12	5	730
Means:	6.92	642.5
NO-LAG GROUP		
1	9	570
2	5	770
3	10	400
4	5	570
5	5	660
6	6	640
7	5	660
8	6	680
9	10	500
10	5	680
11	5	720
Means:	6.45	622.7

time they began learning the characters for lesson one in DeFrancis (3). The no-lag group began learning the lesson one characters at the beginning of the first week. This practice put the lag group about seventy-five characters behind.

The lag group eventually caught up to the no-lag group (in terms of the characters they were responsible for) by the time of the first semester final examination. This was accomplished by having the lag group learn the characters for one additional lesson five weeks into the semester (over the four-day fall break), for another additional lesson eleven weeks into the semester (over the four-day Thanksgiving break), and the third lesson in the fourteenth week (study week), one week prior to the final examination.

In two of the eight weekly class hours, the two groups met together for a grammar lecture, during which time characters were not discussed. In five of the eight class hours, the two

groups met separately with their respective drill instructors. Characters were taught and reviewed during these drill sections.

The two instructors for the two groups are both experienced teachers and native speakers of Mandarin Chinese. Both used only Mandarin as the language of instruction. The groups switched instructors twice (once midway through the fall semester and once midway through the spring semester) in order to help control for the effects of different instructor personalities and methods. Two other instructors taught afternoon sections of the course, and were never involved in teaching either the lag or no-lag group.

*Written Tests.* In the final weekly class hour, the two groups were tested separately. The test instruments administered to each group were identical, except for those portions which tested character writing. The parts of the test which were identical were Part I (listening comprehension), Part II (*pinyin* transcription of unfamiliar Mandarin syllables), and Part III (English-to-Chinese translation, using *pinyin* romanization). The content of Part IV for the no-lag group was character writing (translating English sentences into Chinese characters). The content of Part IV for the lag group, prior to the introduction of characters in the fourth week, was the translation of spoken English sentences directly into *pinyin* Chinese. After characters were introduced to the lag group in the fourth week, Part IV for that group was character writing (the same design as Part IV for the no-lag group, but covering only the characters the lag group had been exposed to).

In the listening comprehension section (Part I) of the weekly quiz, the students listened to Chinese sentences and wrote a translation in English. The audiotaped sentences were presented over headphones in the language lab. The students heard each sentence a total of three times. Upon hearing the Chinese sentences, students were required to do a direct translation from the auditory modality into written English, and were not allowed to write what they heard in *pinyin* as an intermediate step in translating into English.

In the phonetic discrimination portion (Part II), students listened to ten Mandarin monosyllables and were required to correctly transcribe each syllable using *pinyin*, including proper tone marks.<sup>3</sup> For each syllable, one point each was given for a correct initial consonant, vowel (or vowel combination), and

tone. Only phonemes that students had already been introduced to in the course were used in this portion. However, the phonemes were used in novel combinations, thus resulting in unfamiliar syllables. These syllables were audiotaped and presented over headphones in the language lab (immediately following the auditory comprehension portion). Each syllable was heard a total of four times.

In the grammar/*pinyin* section (Part III), students were asked to translate English sentences into Chinese, using *pinyin* orthography. In this section, points were deducted primarily for grammar and vocabulary errors, although minimal credit was lost for *pinyin* (including tone) errors as well.

Part IV (character writing or translation of spoken English) tested different information for the two groups until the final examination. Therefore, a comparison of the two groups for performance on Part IV prior to that point is not included as part of the study.

The fall semester midterm and final examinations contained portions on listening comprehension, grammar/*pinyin*, character reading, and character writing. They did not contain the phonetic discrimination section. Since the character reading and writing portions of the midterm examination contained different information for the two groups (because their character curricula had yet to converge), the comparison of the two groups' performance on these sections is not included in the data analysis. The same final examination was administered to the lag and no-lag groups, since the character curricula of the two groups had converged.

In the second semester both groups took identical tests. Each test (weekly, midterm, and final) contained portions on listening comprehension, grammar/*pinyin*, character reading, and character writing. Listening comprehension and grammar/*pinyin* followed the same design as in the fall semester. Character reading was simply a translation of written Chinese sentences into English, and character writing was a translation of written English sentences into Chinese characters. In the character writing portion, most points deducted were for incorrect characters, although minimal credit was taken off for improper grammar and vocabulary. By design, the sentences in this section were easier than those in the grammar/*pinyin* part.

*Oral Tests.* We administered three oral examinations. The first was given in the seventh week

of the first semester, the second in the twelfth week of the first semester, and the third midway through the second semester. The students were tested in pairs, and were required to ask their partner questions in Mandarin which they translated from English sentences presented to them on slips of paper. Oral production was evaluated based on their performance in asking and answering those questions, and also based on their performance in a short oral interactive session with the instructor.

*Scoring.* The tests were graded as follows. The two parts of the test that required translation into English (i.e., the listening comprehension and reading comprehension sections) were corrected by the same grader every week. This grader is a native speaker of English who is also fluent in Mandarin and had had three years of experience in grading elementary Chinese language tests. The two parts that required translation into Chinese (i.e., the character writing and grammar/pinyin sections) were graded on a rotating basis by the four section drill instructors, all of whom are native speakers of Mandarin.

For the midterm and final examinations, the English portions were graded by the native English grader, and the two Chinese portions were each corrected by one of the Mandarin-speaking drill instructors.

For each of the oral examinations, students' oral proficiency was rated by two instructors as pairs of students asked and answered questions in Chinese and had a brief conversation with the instructor. The students were rated in each of the following five categories: pronunciation, vocabulary, fluency, comprehension, and grammar.

After the tests and examinations were graded, the results were input into a computer by subtest score for data analysis.

#### DATA ANALYSIS

*Areas of Comparison.* How was a comparison of the two groups to be accomplished most effectively? It was possible to compare their performance on any part (or combination of parts) of any individual test, or their performance on any part over a series of tests. For the written tests, it was decided to compare: 1) first-semester final examination scores (since that is where their curricula finally converged); 2) performance over each semester; and 3) performance over the entire year.

For the first semester, performance on the listening comprehension, phonetic discrimination, and grammar/pinyin parts of the weekly tests and the examinations was compared. Performance on the listening comprehension, grammar/pinyin, character reading, and character writing parts of the weekly tests and the examinations was compared over the second semester.

Over the entire year, the only two parts of the written test that were uniformly present were the grammar/pinyin and the listening comprehension. Therefore, the performance of the two groups on only those two written test parts was compared for the entire year.

For oral examinations, the two groups' performance on all three examinations was compared.

*Statistical Design.* For the first semester final examination, a one-way analysis of variance (ANOVA) was performed (with covariates), with group (lag vs. no-lag) as the independent variable, test scores (listening comprehension, grammar/pinyin, character reading, and character writing) as the dependent variables, and VSAT and Chinese language background scores as covariates (to control for the possibility of nonrandom assignment to experimental and control groups).<sup>4</sup>

To analyze test scores over the first semester, a two-way ANOVA (with covariates) was performed, with group (lag vs. no-lag) and instructor as main independent variables; test scores on listening comprehension, grammar/pinyin, and phonetic discrimination as dependent variables; and VSAT and background scores as covariates.

To analyze test scores for the second semester, a two-way ANOVA (with covariates) was performed, with group (lag vs. no-lag) and instructor as main independent variables; test scores on listening comprehension, grammar/pinyin, character reading, and character writing as dependent variables; and VSAT and background scores as covariates.

Test scores over the entire year were analyzed using a two-way ANOVA (with covariates), with group (lag vs. no-lag) and instructor as main independent variables, test scores on listening comprehension and grammar/pinyin as dependent variables, and VSAT and background scores as covariates.

The students' performance on all three oral examinations was analyzed using a two-way ANOVA (with covariates), with group (lag vs.

no-lag) and instructor as main independent variables, the five oral subtest scores (pronunciation, vocabulary, fluency, comprehension, and grammar) as dependent variables, and VSAT and background scores as covariates.

## RESULTS

*First Semester Final Examination.* For the first analysis, the test performance of the two groups was compared at the point in the study when they were first responsible for exactly the same information, viz., the first semester final examination. The mean scores are seen in Table II, and the results of the ANOVA are in Table III.

As may be seen in Tables II and III, the scores of the two groups on the first semester final examination showed no significant differences. Therefore, at the point in the study when the character curricula of the two groups converged, and they were responsible for and tested on precisely the same information, the group which originally was provided a character time lag of three weeks showed no signifi-

cant difference in performance over the group that was given no time lag.

*First Semester Mean Performance.* For the next analysis, the performance of the two groups over the entire first semester was compared. A mean score for listening comprehension, grammar/pinyin, and phonetic discrimination was obtained for each group, as seen in Table IV. For the listening comprehension and grammar/pinyin, data from the ten weekly tests plus the midterm and final examinations (yielding twelve observations per student) were used. For phonetic discrimination, data from the ten weekly tests were used. The results of the ANOVA are in Table V.

As may be seen in Table V, the results of the analysis are that we found no significant main

TABLE II  
Group Mean Scores on First Semester Final Exam: Raw Mean Scores and Mean Scores Adjusted for Constant Score on Covariates\*

Test Section	Group	
	Lag	No-lag
Listening comprehension	20.3	21.0 (raw mean)
	20.0	21.2 (adj. mean)
Grammar/pinyin	21.5	21.5 (raw mean)
	21.5	21.5 (adj. mean)
Character reading	19.9	20.2 (raw mean)
	19.5	20.6 (adj. mean)
Character writing	21.9	22.4 (raw mean)
	21.8	22.5 (adj. mean)

\*Of 25 possible points for each part.

TABLE III  
Summary of Analysis of Variance on First Semester Final Exam Scores

Dependent Variable	Between Groups Sum of Squares	F <sub>(1, 19)</sub> Value
Listening comprehension	8.05	0.78*
Grammar/pinyin	0.01	0.00*
Character reading	5.77	0.34*
Character writing	2.95	0.48*

\*Not significant.

TABLE IV  
First Semester Group Means: Raw Mean Scores and Mean Scores Adjusted for Constant Score on Covariates\*

Test Section	Group	
	Lag <sup>b</sup>	No-lag <sup>c</sup>
Listening comprehension	17.4	17.5 (raw mean)
	17.2	17.7 (adj. mean)
Grammar/pinyin	26.7	26.3 (raw mean)
	26.6	26.3 (adj. mean)
Phonetic discrimination	26.2	25.0 (raw mean)
	26.1	25.1 (adj. mean)

\*For listening comprehension there were 20 possible points, and 30 possible points for both phonetic discrimination and grammar/pinyin.

<sup>b</sup>The total number of observations for the lag group was 136 each for listening comprehension and grammar/pinyin, and 112 for phonetic discrimination. These figures are arrived at as follows. In the first semester, each subject ideally had twelve observations for listening comprehension and grammar/pinyin (ten tests, a midterm, and a final) and ten observations for phonetic discrimination (ten tests). The lag group had twelve subjects, and so for listening comprehension and grammar/pinyin:  $12 \times 12 = 144$ , minus 8 instances of students missing tests = 136. For phonetic discrimination:  $12 \times 10 = 120$ , minus the 8 missing tests = 112.

<sup>c</sup>The total number of observations for the no-lag group was 126 each for listening comprehension and grammar/pinyin, and 104 for phonetic discrimination. These figures are arrived at as follows. In the first semester, each subject ideally had twelve observations for listening comprehension and grammar/pinyin (ten tests, a midterm, and a final) and ten observations for phonetic discrimination (ten tests). The no-lag group had eleven subjects, so for listening comprehension and grammar/pinyin:  $11 \times 12 = 131$ , minus 6 instances of students missing tests = 126. For phonetic discrimination:  $11 \times 10 = 110$ , minus the 6 missing tests = 104.

TABLE V  
Summary of Analysis of Variance on First Semester Test Scores

Dependent Variable	Independent Variable	Between Groups Sum of Squares	F Value
Listening comprehension	Group	11.18	$F_{(1, 257)} = 1.69^*$
	Instructor	0.05	$F_{(1, 257)} = 0.01^*$
Grammar/pinyin	Group	4.36	$F_{(1, 257)} = 0.58^*$
	Instructor	1.81	$F_{(1, 257)} = 0.24^*$
Phonetic discrimination	Group	52.87	$F_{(1, 211)} = 4.83^{**}$
	Instructor	0.27	$F_{(1, 211)} = 0.02^*$

\*Not significant.

\*\* $p < .03$ .

effects for the dependent variables listening comprehension and grammar/pinyin and that, for the dependent variable phonetic discrimination, we found a significant main effect only for the independent variable group. These results mean that over the first semester, there was no difference between groups in listening comprehension or grammar/pinyin, but *there was a significant difference between the groups in phonetic discrimination*. Examining the means of the two groups in Table IV, we see that *the group which was provided the character time lag scored significantly higher than the no-lag group in phonetic discrimination* (26.1 vs. 25.1, respectively). Thus, from these results it would appear that providing students with a time lag results in improved performance in phonetic discrimination, but not in listening comprehension or grammar.

*Second Semester and Entire Year Mean Performance.* The next analyses were comparisons of performance of the two groups over the entire second semester and over the entire academic year. Over the second semester, performance on the listening comprehension, grammar/pinyin, character reading, and character writing parts of the weekly tests and the midterm and final

examinations was compared. Over the entire year, the performance of the two groups on the grammar/pinyin and listening comprehension parts of every test was compared. Results are summarized in Tables VI-IX.

As may be seen in Tables VII and IX, we found no significant differences between groups over the second semester or over the entire academic year. The only significant main effect (as may be seen in Table VII) was for the independent variable instructor, under the dependent variable listening comprehension, during the second semester.<sup>5</sup>

*Oral Examination Mean Performance.* Over the academic year, the subjects had three oral examinations administered, one seven weeks into the first semester, one in the twelfth week of the first semester, and one midway through the second semester. An ANOVA was done on the scores of the students on all three examinations. The results are seen in Tables X and XI.

As may be seen in Table XI, the only significant main effect was for the independent variable group, under the dependent variable fluency. Looking at the means in Table X, we see that the difference was in favor of the group

TABLE VI  
Second Semester Group Means: Raw Mean Scores and Mean Scores Adjusted for Constant Score on Covariates<sup>a</sup>

Group	Listening Comprehension	Grammar/Pinyin	Character Reading	Character Writing
Lag	17.0	26.2	27.7	15.1 (raw mean)
(N = 137) <sup>b</sup>	16.8	26.1	27.5	14.9 (adjusted mean)
No-lag	16.1	26.0	27.2	15.3 (raw mean)
(N = 111)	16.3	26.1	27.4	15.5 (adjusted mean)

<sup>a</sup>For listening comprehension and character writing there were 20 possible points, and there were 30 possible points for both character reading and grammar/pinyin.

<sup>b</sup>The number of observations (N) is arrived at as follows. Over the second semester, each subject ideally had twelve observations (ten tests, a midterm, and a final). The lag group had twelve subjects, and so:  $12 \times 12 = 144$ , minus 7 instances of students missing tests = 137. The no-lag group had ten subjects (one group-two subject's data were unavailable for analysis of variables involving second semester), and so:  $10 \times 12 = 120$ , minus 9 instances of students missing tests = 111.

TABLE VII  
Summary of Analysis of Variance on Second Semester Test Scores

Dependent Variable	Independent Variable	Between Groups Sum of Squares	F <sub>(1, 243)</sub> Value
Listening comprehension	Group	11.18	0.98*
	Instructor	47.61	4.18**
Grammar/pinyin	Group	0.05	0.01*
	Instructor	0.06	0.01*
Character reading	Group	0.37	0.05*
	Instructor	0.08	0.01*
Character writing	Group	14.90	1.56*
	Instructor	16.15	1.69*

\*Not significant.

\*\*p < .05.

which had been provided the time lag (lag = 19.3 vs. no-lag = 18.6). This means that, over all three oral tests, *the group which received the character time lag was judged to be significantly more fluent than the no-lag group in their oral production of Mandarin Chinese.*

#### SUMMARY AND DISCUSSION

The students who were provided with a three-week time lag prior to the introduction of characters into their elementary Chinese curriculum were better able to discriminate phonetically and transcribe unfamiliar Mandarin syllables — and were also more fluent in spoken Mandarin — than the students who were not provided such a lag. The no-lag group was not consistently found to be significantly better than

TABLE VIII  
Academic Year Group Means: Raw Mean Scores and Mean Scores Adjusted for Constant Score on Covariates<sup>a</sup>

Group	Listening	
	Comprehension	Grammar/Pinyin
Lag (N = 273) <sup>b</sup>	17.2	26.4 (raw mean)
No-lag (N = 227) <sup>b</sup>	17.0	26.3 (adjusted mean)
	16.8	26.1 (raw mean)
	16.9	26.2 (adjusted mean)

<sup>a</sup>Of 20 possible points for listening comprehension and 30 possible points for grammar/pinyin.

<sup>b</sup>The number of observations (N) is arrived at as follows. Over the entire year, each subject ideally had twenty-four observations (twenty tests, two midterms, and two finals). The lag group had twelve subjects, and so: 12 × 24 = 288, minus 15 instances of students missing tests = 273. The no-lag group had ten subjects, and so: 10 × 24 = 240, minus 13 instances of students missing tests = 227.

TABLE IX  
Summary of Analysis of Variance on Academic Year Test Scores

Dependent Variable	Independent Variable	Between Groups Sum of Squares	F <sub>(1, 495)</sub> Value
Listening comprehension	Group	1.36	0.15*
	Instructor	25.81	2.79*
Grammar/pinyin	Group	3.56	0.44*
	Instructor	0.48	1.69*

\*Not significant.

TABLE X  
Group Means for Three Oral Exams: Raw Mean Scores and Mean Scores Adjusted for Constant Score on Covariates<sup>a</sup>

Oral Test Score	Group	
	Lag (N = 36)	No-lag (N = 30) <sup>b</sup>
Pronunciation	17.9	17.9 (raw mean)
	17.9	18.0 (adj. mean)
Vocabulary	18.1	18.0 (raw mean)
	17.9	19.1 (adj. mean)
Fluency	19.4	18.5 (raw mean)
	19.3	18.6 (adj. mean)
Comprehension	19.5	19.4 (raw mean)
	19.5	19.5 (adj. mean)
Grammar	17.6	17.7 (raw mean)
	17.8	17.8 (adj. mean)

<sup>a</sup>Each of a possible 20 points.

<sup>b</sup>The N of 36 represents 3 oral tests times 12 subjects in the lag group, and the N of 30 represents 3 times the 10 subjects (who participated over both semesters) in the no-lag group.

TABLE XI  
Summary of Analysis of Variance on Oral Exam  
Test Scores

Dependent Variable	Independent Variable	Between Groups	
		Sum of Squares	F <sub>(1, 61)</sub> Value
Pronunciation	Group	0.10	0.04*
	Instructor	0.32	0.13*
Vocabulary	Group	0.37	0.13*
	Instructor	0.21	0.07*
Fluency	Group	7.43	5.18**
	Instructor	3.10	2.16*
Comprehension	Group	0.02	0.02*
	Instructor	0.10	0.14*
Grammar	Group	1.11	0.39*
	Instructor	1.37	0.48*

\*Not significant.

\*\*p < .03.

the lag group in any aspect of Mandarin Chinese that was analyzed as part of this study.<sup>6</sup> The use of VSAT and Chinese language background scores as statistical controls provides assurance that the observed effects were not due to differences between subjects in verbal ability or amount of prior Chinese language background.<sup>7</sup>

These results support the idea that providing students with a grace period before characters are introduced into the elementary Chinese curriculum is beneficial to the second language acquisition of Chinese. They suggest that not providing such a grace period may detract both from students' ability to discriminate Mandarin sounds and from their fluency in the spoken language.

Furthermore, since no consistent differences were found between the two groups in character reading or writing, this indicates that it is spoken language, and not literacy, that benefits from the delay in the introduction of character orthography.

To explain the decreased performance in phonetic discrimination, I suggest that the

initial focus on nonphonetic character orthography results in a reduced awareness of the sound structure of the language. Hayes (pp. 191-92) found that, contrary to native readers of Chinese, second language students of Chinese do not use a primarily phonological strategy in reading characters. If his findings are true, then students who initially concentrate on characters could have a reduced awareness of the phonological shape of the language, which would explain their decreased performance in phonetic discrimination. These results support Hayes' contention (p. 194) that if students initially forgo learning character orthography and first build a foundation in the spoken language, they may develop superior phonological inferencing skills.

As for the reduction in oral fluency, we may hypothesize that the initial focus on characters resulted in a reduced overall facility with the spoken language, which was manifest in a lower fluency score.<sup>8</sup> Considering Hayes' results once again, the eschewal of a primarily phonological strategy by second language students of Chinese means that knowledge of character orthography for these students is to a certain degree "removed" from the production of spoken language. Under these conditions, the knowledge of character orthography may act to inhibit rather than facilitate the production of oral speech, thus resulting in a lower degree of fluency.

The present study demonstrates that providing students with a time lag prior to the introduction of an unfamiliar, nonphonetic orthography in a second language curriculum yields measurable benefit and no apparent harm in the acquisition of the target language. At this point, it is unclear whether it was the non-familiar rather than the nonphonetic nature of the orthography which caused the difference between the two groups. Providing a time lag for a language with an orthography which is unfamiliar but phonetically regular (e.g., Russian, Arabic, or Japanese) would possibly yield little or no benefit in the acquisition of that language.

#### NOTES

<sup>1</sup>This study was supported by a grant from the Consortium for Language Teaching and Learning. I am grateful to the Consortium and its executive director, Peter Patrikis,

for this support. I would also like to thank an anonymous *MLJ* reviewer for helpful comments.

<sup>2</sup>Our experience indicates that sections which are scheduled at different times tend to attract students with considerable differences in motivation, course goals, and ability.

<sup>3</sup>Mandarin Chinese has four tones (high level, mid rising,



low dipping, and high falling) which are used in a phonologically contrastive manner, in the same way as consonants and vowels.

<sup>4</sup>Covariates are measures taken of pre-existing characteristics of subjects (such as, in our case, verbal ability or amount of background in the target language) which may affect subjects' test performance scores. These measures are introduced into the analysis of variance (more properly, analysis of covariance) in order to ensure that observed differences between the experimental and control groups on test scores are not due to differences in those pre-existing characteristics. For a more complete discussion of the use of covariates, including procedures for determining degrees of freedom, see Winer (chapter 10).

<sup>5</sup>This result indicates that over the second semester, the students taught by one of the instructors scored significantly higher in listening comprehension than the students taught by the second instructor (covariate adjusted means for the two instructors: 17.0 vs. 16.1;  $F_{(1, 243)} = 4.18$ ,  $p < .05$ ).

<sup>6</sup>There were a few instances where one of the groups performed significantly better than the other on a given part

of an individual *weekly* test. For example, the no-lag group performed better on the grammar/pinyin part of the fourth test of the first semester and the third test of the second semester, while the lag group performed better on the grammar/pinyin part of the fifth test of the second semester. We attribute results such as these either to one-time effects or to random effects that were not related to the time lag.

<sup>7</sup>Many effects were found due to these covariates, with the VSAT found to be a better predictor of student performance than background score. In general, the higher the VSAT score, the better the students performed on grammar/pinyin, listening comprehension, phonetic discrimination, and character reading and writing, and the higher the background score, the better the students performed on listening comprehension, character reading, and the "vocabulary" subpart of the oral examinations.

<sup>8</sup>The results on oral fluency are statistically robust. That is, not only was the effect for fluency the only one found to be significant for the three oral tests combined, but it was also the only effect that was significant for *each* of the three oral tests.

---

#### BIBLIOGRAPHY

---

1. Chang, Richard. *Read Chinese: Book II*. New Haven, CT: Far Eastern, 1983.
2. DeFrancis, John. *Beginning Chinese*. New Haven, CT: Yale Univ. Press, 1976.
3. ———. *Character Text for Beginning Chinese*. New Haven, CT: Yale Univ. Press, 1976.
4. Hayes, Edmund. "Encoding Strategies Used by Native and Non-Native Readers of Chinese Mandarin." *Modern Language Journal* 72 (1988): 188-95.
5. Walker, Galal. "'Literacy' and 'Reading' in a Chinese Language Program." *Journal of the Chinese Language Teachers Association* 19 (1984): 67-84.
6. Winer, B. *Statistical Principles in Experiment Design*. New York: McGraw-Hill, 1971.