

THE EFFECT OF DELAYED COMPARISON IN THE LANGUAGE LABORATORY ON PHONEME DISCRIMINATION AND PRONUNCIATION ACCURACY

Cyrus R. Sisson
The University of Michigan

Three separate classes of students of English as a second language were divided into groups matched on measures of pronunciation discrimination and production. One half of the students monitored their previously recorded responses to pronunciation exercises in the language laboratory and the other half received comparable exposure to the contents of these exercises by responding actively a second time. A mobile language laboratory, The Plurilingua, used in this experiment, which permits up to six individuals to record and subsequently monitor their responses on one spool of magnetic tape is described. It was hypothesized that the active groups would show a significantly higher degree of improvement in discrimination accuracy and pronunciation authenticity of English phonemes. At the end of the experiment, no significant effects were found, although there was evidence of a trend which favored the active groups.

One of the most prolonged controversies that has developed since the adaptation of the tape recorder as an instrument of second language pedagogy, and one that has important instructional and economic ramifications, is whether or not standard language laboratory procedures should include a delayed comparison cycle in which students individually record their responses to exercises and then monitor them passively.¹ It would seem to be a relatively simple matter to evaluate this question scientifically and to reach a decision supported by incontrovertible evidence. Lamentably, such has not been the case, and the language laboratory literature is replete with examples of argument and counter-argument proffered by the partisans of the opposing schools of thought on this matter. Notoriously absent from the literature are carefully controlled studies aimed at the elucidation of this problem. To be sure, so-called experiments have been reported and articles have been published in support of both points of view. Generally speaking, however, the articles cite no evidence and the studies are egregiously lacking in control with the result that they are seldom applicable or replicable.

¹This research was supported in part by The English Language Institute, The University of Michigan, and by the U.S. Office of Education, Contract OEC-0-9-097740-3743.

In view of the paucity of worthwhile experimentation dealing with this question, the dogmatism with which educators have expounded their particular biases is surprising. In discussing this aspect of language laboratory planning, Watkins remarked that "Many instructors share our opinion that in a beginning course, though it may be chiefly based on pattern assimilation as it is here, *recording* is actually of little real value." (1960: 63) Adopting a somewhat different line of reasoning, Morton (1961), while recognizing that opinion on the matter was divided, stated that feedback or reinforcement must be accomplished within approximately three seconds of initial response to be optimally effective. His conclusion therefore was that delayed feedback as provided by conventional tape recorders is hardly useful. On the other hand, a discussion of the same question at the Northeast Conference on the Teaching of Foreign Languages recommended incorporation of the delayed comparison cycle, which, they noted, ". . . will allow the teacher to locate and point out to the student habitual mistakes, and will lead him to perceive and to correct them and to understand better the structural characteristics of his own language and, in contrast, of the target language." (Bottiglia 1957: 62) This same Conference though, adopted a somewhat modified position for colleges and universities several years later.

At present most colleges do not have enough competent teachers to provide constant instructional supervision in the various languages. There is no machine now known that can perform one of the most important roles of the instructor, to detect, analyze, and correct mistakes in pronunciation. A student who fails to give the morphologically and structurally correct response to a stimulus will, in most cases, be able to recognize his error when the tape gives him the model response right after his own performance. Mere phonetic differences from the model, on the other hand, are much less easily detected and corrected by the student on his own. If a competent instructor cannot be provided for this purpose, the student may waste a good deal of time in recording his faulty pronunciation and listening complacently to it, with the danger of fixing his errors instead of correcting them.

Although student recording under competent individual supervision is probably the best way of correcting pronunciation, we may conclude that whenever such competent assistance cannot be provided in the college laboratory, the student should not be allowed to spend a great deal of time recording and listening to his own voice. Instead he should spend his time listening to good models and imitating them with or without the help of activated earphones. (Flaxman 1961: 58)

The same situation was referred to by Lado in the following way:

It has been argued that the student should not listen to his own mistakes but should listen only to the master recording. This is hardly relevant, since the student listens to his mistakes in the audio-active laboratory anyway. It is true, however, that there is no point in recording the student's responses when he feels he is still struggling to improve them and knows what he is doing wrong. When the student has progressed as far as he is able in discernment and thinks he is right, the full value of the dual-channel recording equipment is reached. At that point, he should record his responses and compare them objectively with the master. He will detect problems he is not able to detect otherwise. And when the teacher points out his mistakes to him, he can concentrate on observing them on the tape without having to devote his attention to the complex operation of speaking the foreign language. (1964: 186)

Many recent articles dealing with second language learning are also quite authoritative in their recommendations regarding the recording language laboratory. (Hamilton 1970: 477-8)

The laboratory also should be used by the individual student in making his own tapes. The opportunity to hear his own errors is a great advantage to the student unconscious of them. These tapes are also useful to the instructor who, try as he might, is hard put to note every error he hears: a permanent taped record is invaluable for the instructor's further study and subsequent review with the student. The student should also keep all his recorded assignments for repeated review.

Perhaps one of the most adamant advocates of the recording laboratory is William N. Locke. He has inveighed against the proposal of some authors to substitute the audio-active laboratory (in which students hear themselves as they speak, but neither record nor subsequently monitor their responses), for the recording laboratory and the delayed comparison cycle. Basically his argument hinges on two main points. "Bone conduction affects every sound heard by the student as he speaks. It affects his evaluation of those sounds, applying equally to his own and foreign languages." (1960: 278) The reason we can never "hear ourselves as others hear us except through recording," is that only the airborne component of speech is recorded. Locke feels that the only way an individual can compare his pronunciation, intonation, and accent with those of another speaker is by objectively listening to his recorded speech wherein that component of the speech signal distorted by bone conduction is eliminated. His second major point is concerned with a psychological aspect of the question. "Has a student a sufficient capacity for attention to (1) remember, while (2) he speaks, while (3) he hears, while (4) he compares and judges

himself? If the answer is 'yes,' then it is less necessary for him to record his voice; if it is 'no,' then he must record." (1960: 278)

Mueller considers the same phenomenon under a somewhat different light. He posits an essential interplay between speech organs and sound perception, and considers it essential in the acquisition of new speech habits. He consequently sees a deleterious effect in the incorporation of a delayed comparison cycle in standard language laboratory procedures.

Delayed hearing, that is, hearing one's performance at the end of the exercise by playing the tape back, does not involve the total organism. When the student records, he cannot at that moment check the sounds he is making. Since we are slaves of habit, he is likely to use his speech organs according to the pattern of his native tongue, thus not drilling the new speech habits. When he hears his performance and compares it with the model at the conclusion of the exercise, his speech organs are no longer involved. He registers the discrepancy mentally, rather than correcting automatically the position of his tongue and lips. It can be compared to a person learning to use a tool with his hands, but he is blindfolded and is permitted to see the results only at the end of every try. He cannot see how he uses his muscles and the mistakes which result from such use; he therefore cannot instantly correct mishandling of his tool. (1958: 171)

His conclusion then is that the student should hear his responses instantaneously when he repeats the foreign speech. The interplay between muscles and auditory perception should not be separated and drilled individually.

Additional factors have been pointed out, however, which may have equal relevance to the delayed comparison controversy. Hayes maintains that "Facilities permitting the student to record his own voice and compare it with a model seem to increase student motivation. He enjoys the whole procedure, and certainly any feature that increases motivation should be considered valuable." (1963: 38) Mathieu, however, considers the delayed comparison cycle to have a diametrically opposite effect on student motivation. "In general . . . AEP [audio-evaluatory practice] has drawbacks. If a student is asked to record and listen back to patterns that are new to him, his pronunciation and fluency will be quite imperfect and the AEP will undermine his self-confidence instead of increasing it." (1960: 123)

Other authors reject the recording language laboratory for administrative as well as pedagogical reasons. Kieser offered the following commentary:

In dealing with the language laboratory we are going to ignore the laboratory with full play-back, because we have al-

ways been of the opinion that play-back has no real place in secondary schools, except perhaps at the grade 13 level. No doubt senior high school students and university students can derive benefit from the comparison of model speech with their own efforts, but to subject the average grade 9 student to the play-back of a completed pattern drill is not only a waste of time but is courting disaster. It is of paramount importance to counteract by all possible means, every opportunity that a student may have for inaction or merely passive listening, because any knobs, buttons or switches are fair game for idle hands and an unoccupied mind. (1964: 32-3)

The entire controversy was epitomized by Hocking who concluded that the argument was no longer meaningful in any event.

This last technique, also called "record-playback-compare," was generally considered to be a unique and highly valuable experience. After a time, however, this assumption was challenged. . . . The chief objection concerned the alleged reinforcement of error by hearing it played back, and it was claimed that the time required to play back would be better devoted to additional practice in hearing and responding without recording, yet simultaneously hearing the response through the activated microphone and headset. The counter-argument was that so many simultaneous activities were unsatisfactory, that bone conduction prevented objective hearing, and that recording provided the only truly objective comparison for self-criticism.

The most recent thinking gives little attention to either imitation or recording as such and more attention to the discrimination and reproduction of fine distinctions of sound. Intensive training exercises in the audio discrimination of two or three sounds are to be associated with similar exercises in oral production, both of them fortified by brief recordings with immediate playback. Thus the student always has available an *immediate* playback of limited duration, in addition to the extended playback provided by rewinding. The immediacy of playback heightens attention just as the concentration of the exercises on one or two fine distinctions also focuses attention; in short, refinement of method combined with refinement of equipment.

The old argument over record-and-playback has thus been superceded by sophisticated considerations of perception, discrimination, and reproduction. It is no longer thought that beginning students can profit from playback of unselected materials, just as it is no longer assumed that sheer quantity of oral imitation is necessarily a virtue. (1964: 31)

In spite of Hocking's dismissal of the whole controversy, a dismissal itself unsupported by empirical evidence, the question remains relevant. Language laboratories offering recording and delayed comparison capabilities abound, as do recorded language teaching materials which can be employed with or without delayed comparison. Since the opinions of educators about the optimum use

of this teaching aid seem divided, the logical source for resolution of the problem would seem to be among the published experimental studies into the effectiveness of the language laboratory in general and the delayed comparison cycle in particular.

Unfortunately, only three serious studies exist which deal in any substantial way with the advantages and disadvantages of a particular kind of laboratory procedure. Those that have been published include the Keating Report (Keating 1963), the Pennsylvania Project (Smith 1968), and the series of studies conducted in New York City high schools by Sarah Lorge and Edward G. Bernard (1959; 1962; 1963; 1964; 1965). The attacks on the methods employed in the Keating Report have been both numerous and cogent (Anderson 1964; Edling 1963; Gaarder & Hutchinson 1963; Grittner 1964; Mathieu 1963; Porter & Porter 1964; Stack 1963; 1964). Similarly, the Pennsylvania Project has been criticized for its failure to control for teacher ability, motivation, and experience; its use of tests of dubious applicability; and its assignment of teachers to teaching strategies not of their preference (Aleamoni 1969; Clark 1969; Hocking 1969; Otto 1969; Valette 1969). In his criticism of the Pennsylvania Project, Hocking remarked that "... involved too many imponderables and uncontrolled variables." His recommendation was that "Such research projects should be abandoned in favor of the more restricted projects in which the major variables can be controlled." (1969: 410)

Of the three, only the Lorge-Bernard studies systematically delved into performance differences among students whose language laboratory exposure differed in terms of the presence or absence of a delayed comparison cycle (Lorge 1964). Among other findings, it was reported in this study that insofar as overall quality of speech was concerned, the recording-playback daily attendance group made significantly greater gains at the .01 significance level than either the other laboratory groups or the non-laboratory group. No significant over-all pattern of difference between audio-active and record-playback groups emerged.

Although most reviews of the Lorge-Bernard findings have been favorable, serious shortcomings have been indicated by J. R. Green (1965), especially lack of uniformity in the language learning environment, because the classes were taught by different instructors.

Thus, in the literature on language laboratory methods, the question of the advisability of the delayed comparison cycle has elicited a polarity of opinion and a corresponding range of recommendations. Few relevant studies have been published, and those that have been presented have been subjected to severe criticism for their failure to control for important variables affecting second

language learning. The value of the recording and the subsequent monitoring of responses is undetermined, but, in view of the existence and continuing commercial sale of both recorded language teaching materials and laboratory installations incorporating the delayed comparison cycle, it is worth investigating.

An inexpensive, mobile language laboratory with a delayed comparison cycle capability was used in the present study of the effect on pronunciation discrimination and authenticity of two different modes of language laboratory use. The experiment was limited to the effect of a delayed comparison cycle on the perception and production of English phonemes by students of English as a second language. No attempt was made to assess the effectiveness of a delayed comparison cycle in the teaching of grammatical constructions or vocabulary.

It has been noted that many writers in the field maintain that students in the language laboratory either do not recognize their own errors during the playback cycle or else allow their minds to wander and hence do not attend to material that does not require active participation. It was hypothesized that, *ceteris paribus*, if matched groups of students are exposed to identical conditions in both the classroom and the language laboratory except that half listen passively to each pronunciation exercise immediately after recording it (delayed comparison), while the other half respond to the same exercise a second time, the active group will show a significantly higher degree of improvement in both the discrimination and the production of English phonemes at the end of the experiment.

METHOD

Equipment

A self-contained recording language laboratory, conceived by Emmanuel Companys, was constructed by Timothy Rand in the laboratories of the Center for Research on Language and Language Behavior. This laboratory, the Plurilingua, has the following characteristics. Six student positions are controlled by a monitor from a four-track Viking master tape-deck, model 230-1900 with amplifiers RP 110 and RP 120. This system operates with a single seven inch reel of standard magnetic tape, containing a permanently recorded lesson to which the responses of a maximum of six students can be temporarily or permanently recorded simultaneously. The lesson tape itself must be specially processed initially to provide this capability.

Consider a lesson of the "cascade" type in which either the

subject or the predicate may be changed in the basic model sentence "John slept."

Col 1	Col 2	Col 3	Col 4	
(Stimulus)	(Response)	(Confirmation)	(Repetition)	
Mary	Mary slept	Mary slept	Mary slept	
ate	Mary ate	Mary ate	Mary ate	
walked	Mary walked	Mary walked	Mary walked	
Peter	Peter walked	Peter walked	Peter walked	etc.

The pre-recorded parts of common commercial recordings consist of columns 1 and 3 with suitable silent intervals left between them during which the student responds. The Plurilingua tapes were prepared in a somewhat different manner to allow six students to use one tape deck in concert and still be able to record their responses individually and subsequently monitor them when a delayed-comparison cycle is deemed desirable. All language laboratory lessons were re-recorded from the original script by both male and female native speakers of English under the supervision of the author. The stimuli and confirmations were recorded on the top channel but instead of leaving an interval between them, both the stimulus and the confirmation were repeated a second time in place of the silent interval which normally follows them. The pre-recorded model might then be represented something like the following:

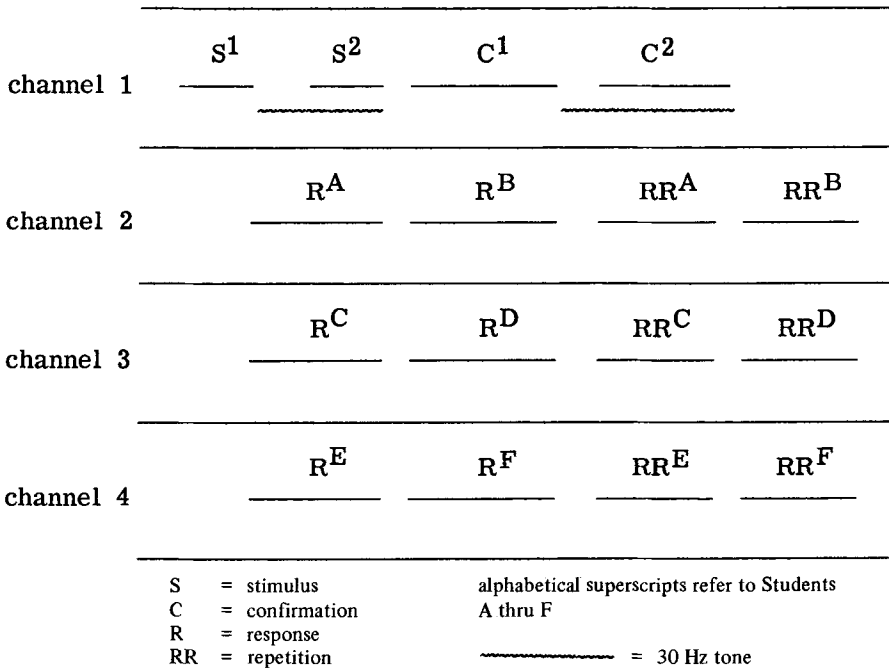
S ¹	S ²	S ¹	S ²	S ¹	S ²	
Mary	Mary	Mary slept	Mary slept	ate	ate	etc.

Half the six students hear the first set of stimuli (labelled S¹) and the other half hear the second set (labelled S²). While the second half of the students are listening to S² (stimulus or confirmation), the first half are responding, one per track; similarly, while the second half of the students are responding, the first half are listening to the next stimulus.

In the preparation of the top channel an inaudible 30 Hz tone is superimposed during the second of each element of a stimulus/confirmation pair. A voice-operated relay (VOR) and filter combination is employed to switch the pairs of students in any one channel from the listening mode (top channel) to the recording mode (one of the bottom three channels). Similarly in a delayed-

PHONEME DISCRIMINATION--PRONUNCIATION ACCURACY 77

comparison cycle, each student hears the appropriate element from the top channel and only his own response from one of the bottom channels. Sufficient time is left between the two recordings of any stimulus/confirmation pair to permit the student to complete responses significantly longer than the stimulus. Thus, if the responses of students A and B are recorded on channel 2, of students C and D on channel 3, and of students E and F on channel 4, a recorded tape would look schematically like this:



The listen-response pattern of a single student position, that of student D, for example, operates in the following manner: during the presentation of S¹, student D is technically in the record position, but not having as yet been presented with a stimulus, he does nothing. With the detection of the 30 Hz signal by the filter-VOR combination, the Plurilingua switches position D to the listening mode and student D in concert with students B and F, listens to S². Once the 30 Hz tone ends, position D is placed in the record mode which is signalled to the student by the illumination of a small individual light. Student D's response is recorded on channel 3 in the place indicated in the schematic diagram above (R^D). When the Plurilingua detects the onset of the following 30 Hz signal, the

student signal-light is extinguished, position D is placed in the listen mode and he hears the confirmation C². At the end of the confirmation, the position is switched to the response mode (again signalled by the illumination of the position light) and the student repeats the confirmation (RR^D) which is recorded on channel 3. The sequence continues until the end of the exercise, at which time the monitor can present the exercise actively again, playback the recorded responses (delayed-comparison), or continue on to the following exercise.

One advantage offered by the method of recording two sets of student responses per channel is that by making use of another of the options provided by the Plurilingua, the instructor can monitor two students in the amount of time normally spent monitoring the recording of one student. He simply elects not to listen to the model on channel 1, and instead monitors the responses of both on any given channel.

It should be obvious from the foregoing that because of the nature of the machine's operation, the number of elements in each unit of the lesson must be even. In the event it is deemed pedagogically desirable for the student to hear a confirmation, but not to repeat it, the second stimulus set would be recorded directly after the first confirmation set, by pairs, without a pause in the 30 Hz signal. Thus the recording of the same exercise would look like this:

<u>s¹</u>	<u>s²</u>	<u>C¹</u>	<u>C²</u>	
Mary	<u>Mary</u>	Mary slept-ate	<u>Mary slept-ate</u>	
	<u>R^A</u>	<u>R^B</u>	<u>RR^A</u>	<u>RR^B</u>
	Mary slept	Mary slept	Mary ate	Mary ate
	<u>R^C</u>	<u>R^D</u>	<u>RR^C</u>	<u>RR^D</u>
	Mary slept	Mary slept	Mary ate	Mary ate
	<u>R^E</u>	<u>R^F</u>	<u>RR^E</u>	<u>RR^E</u>
	Mary slept	Mary slept	Mary ate	Mary ate

In the original model of the machine, the tape deck was augmented by a control panel which allowed the operator to place all positions in the listening mode and to connect them to channel 1 simultaneously so that only one set of instructions and examples had to be prepared. The tape deck contains an electric eye operated circuit which halts operation of the machine when the tape is exhausted. This feature was adapted to permit automatic stopping of the Plurilingua between instructions and exercise, by splicing a six-second segment of translucent leader between them and at the end of each exercise before the following instructions. This particular strategy, while effective, was inefficient for three reasons. It was time-consuming to prepare the tapes by hand; the tapes once spliced could never be re-used for another lesson with a different number of exercises, or exercises of a different length; and finally, the tapes could not be machine-duplicated, thus eliminating to a large extent the possibility of successful commercial use of the machine and its tapes.

The problem was solved in a manner analogous to that employed to switch the students from the listening mode to the record mode; i.e., the addition of yet another tone. An additional VOR and a band pass filter were installed in the experimental model of the Plurilingua. All of the taped lessons were re-recorded with the addition of a 50 Hz tone which did not register wherever there was leader on the original tape. When these altered lessons were then copied on fresh, uncut magnetic tape, the copies had sections between lessons and instructions on which there was no signal. During their playing, the original set of 30 Hz tones triggered the same functions as before, while the absence of the 50 Hz tone caused the machine to stop, at which time the operator could make the changes necessary to switch the machine from the listening mode to the record mode. Lessons processed in this fashion were easier to manufacture originally, did not prevent the magnetic tapes on which they were recorded from being re-used, and could be duplicated automatically.

The final innovation added to the Plurilingua after its first period of testing was a switch which permitted the operator to select one of two methods of utilization. The machine was adapted to proceed through an entire seven inch reel in one of the two following modal sequences.

- 1) Instructions exercise 1, record exercise 1,
rewind to beginning of exercise 1,
re-record exercise 1,
Instructions exercise 2, record exercise 2, rewind,
re-record exercise 2

- 2) Instructions exercise 1, record exercise 1, rewind to beginning of exercise 1, listen to individual recording of exercise 1,
Instructions exercise 2, etc.
to the end of the tape

The second sequence provides the option of a delayed-comparison cycle, while the first actively presents the same exercise sequence in the same order and for the same elapsed amount of time as in the delayed-comparison cycle.

Basically then, the laboratory is a standard four track tape recorder which has been modified by the addition of extra amplifiers, filters, voice-operated relays, and other relays to permit the responses of six students to be recorded during a single laboratory session on one spool of specially prepared standard magnetic tape. These tapes may be used repeatedly by successive groups of students without intermediate processing, or the responses of a particular class of six may be retained as a permanent record.

Subjects

At the English Language Institute, The University of Michigan, incoming students are tested upon their arrival to determine the degree to which they have already acquired English language skills. Each student receives a score in grammar, auditory comprehension, and vocabulary. On the basis of these scores, students are divided into relatively uniform sections with respect to displayed knowledge of English. The students of the intensive eight weeks' course with the lowest scores, generally between 9 and 12 students, are placed in section 81. They spend four hours a day in the classroom studying grammar, pronunciation, vocabulary, and finally the incorporation of all these skills in a class called pattern practice. The students attend their classes together at the same time every day, and for any given eight-week period, each class is always taught by the same four teachers. In addition to these four class periods, the normally enrolled ELI students spend an additional hour in an audio-active language laboratory without recording facilities, where they listen and respond to four series of tape recordings corresponding to the four class types. The typical exercise for grammar and pattern practice consists of three elements: an auditory stimulus (often accompanied in the case of pattern practice lessons by a drawing), the student response, and a confirmation of the response followed immediately by the next stimulus. No time is provided for repetition of the confirmation. The pronunciation and vocabulary lessons generally consist of simple

repetitions, of the form S-R, without confirmation. In addition, on the pronunciation tapes there are short quizzes in which the students hear words selected at random from two sets of minimal pairs.

Column one	Column two
sail	sell
mate	met
bait	bet
laid	led
wait	wet
late	let

Each time he has to indicate whether the word comes from column 1 or column 2, a fairly straight-forward identification task. Finally, large sections of the vocabulary tapes are simply dialogues that the students listen to passively.

Procedure

The test of Aural Perception for Latin-American Students developed by Robert Lado to investigate English phoneme discrimination by non-native speakers of English was modified by the experimenter to form the basis of a three part discrimination-production test used to match Ss. Part I of the modified examination—the discrimination portion—consisted of four sections, the first two involving sets of words which differed by one phone and the last two involving sets of sentences which differed by one phone. In subparts I and III, Ss decided which of three elements, words or sentences, were alike, and marked their responses on a test form. In subparts II and IV, a model was presented followed by three test items to be compared with it. Ss were told that none, one, two, or all three of the test items could be the same as the model word or sentence and again instructed to mark their responses on the test form. In the second major part of the test, Ss mimicked a single word presented once auditorily. Prior to part III, a list of 50 words was read aloud to Ss who were provided with a printed version of the list. In part III, each S read the same list of words aloud, one at a time, in a different random order, when signalled. The test and all instructions, together with several examples requiring active participation, were pre-recorded and presented using the Plurilingua. The instructions and examples were played again

in case of questions or misunderstanding. All oral responses were recorded and subsequently played back to judges, native speakers of English, ignorant of the correct response, who transcribed only the single phoneme being tested using an answer sheet which provided the surrounding phonemic environment. Initially, two judges, S. Josephson and the present author, evaluated each response independently, with perfect agreement. This remarkable interjudge reliability is understandable in light of the fact that after transcription, an S's score was the total number correct. A badly pronounced phoneme identified differently by each judge was nevertheless incorrect. After testing the first of the three groups of Ss, only one judge was thereafter employed. For each transcription, however, the judge was ignorant of the desired oral response and simply transcribed the test phoneme he heard.

After initial testing, each class was divided into subsections for the purpose of language laboratory attendance only. All other classes were attended together. During the first week of the course, one subsection attended mandatory laboratory sessions for two hours in the morning and the other for two hours in the afternoon under the supervision of the experimenter. Every subsequent week, this schedule was reversed so that by the end of the course each subsection had attended the same number of morning and afternoon obligatory laboratory sessions, evenly distributed throughout the term. Thus, the variables of learning environment were controlled as closely as possible with respect to identity of instructors, scheduling of laboratory lessons, and use of classroom and laboratory materials.

All subsections spent exactly the same amount of time in the laboratory and the single difference in treatment involved the manner in which the second exposure to each pronunciation exercise transpired. Members of one subsection listened passively to their pre-recorded individual responses to the exercise (delayed-comparison cycle) and the other subsection responded actively to the exercise the second time.

RESULTS

The scores obtained by each student on the pre-test used to match the groups and on the identical post-test are presented in Tables I, II, and III, together with each S's age and his score on the English Language Institute's admission tests (grammar-G, auditory comprehension-AC, and vocabulary-V). A statistical analysis of these data reveals that there are no significant differences between the groups. Furthermore, there are not even any consistent

TABLE I

Individual and group mean scores for both active and delayed-comparison (DC) groups for the English Language Institute Placement Test, and both the pre-tests and post-tests in phoneme discrimination, mimicry, and oral reading. Column I indicates the age of each S.

	PLACEMENT TEST		DISCRIMINATION TEST				PRODUCTION TESTS		
	I	II	III	IV	Mimicry	Reading	δ	δ	δ
Group I-A Active									
1. Subject A	17	70 54 66	72 80	40 60	68 80	40 60	56 80	54 70	
2. Subject B	35	76 58 74	68 80	48 76	84 92	56 80	66 70	67 74	
3. Subject C	22	80 74 76	84 96	56 64	76 88	56 84	56 84	70 82	
4. Subject D	29	46 40 46	44 88	28 44	72 68	40 56	48 66	44 70	
\bar{X}		68 56½ 65½	67 86	43 61	75 82	48 70	16 56½ 75	18½ 58 74	15½
Group I-A DC									
1. Subject E	31	74 56 68	64 56	40 56	56 48	40 48	54 62	50 58	
2. Subject F	22	62 40 36	84 84	56 68	64 84	64 72	80 94	52 56	
3. Subject G	25	84 68 72	92 96	24 60	56 68	36 64	54 70	52 68	
4. Subject H	17	92 60 80	76 92	56 64	64 76	40 60	49 78	60 72	
\bar{X}		78 56 64	79 82	44 62	60 69	45 61	11½ 59½ 76	16½ 53½ 63½	10

TABLE II

Individual and group mean scores for both active and delayed-comparison (DC) groups for the English Language Institute Placement Test, and both the pre-tests and post-tests in phoneme discrimination, mimicry, and oral reading. Column 1 indicates the age of each S.

	PLACEMENT TEST	DISCRIMINATION TEST								PRODUCTION TESTS			
		I	II	III	IV	Mimicry	Reading						
Age	G AC V	P O R S E T	P O R S E T	P O R S E T	P O R S E T	P O R S E T	P O R S E T	δ	δ	δ	δ	δ	
Group II-A Active													
1. Subject I	26 86 70 66	48 88	48 76	48 60	52 64	72 82	64 72						
2. Subject J	29 66 62 74	84 84	52 68	68 80	52 80	70 76	52 66						
3. Subject K	43 92 74 88	44 92	52 64	60 76	56 72	70 74	76 78						
4. Subject L	28 86 78 80	72 96	68 72	72 68	64 72	74 84	58 64						
\bar{X}	$82\frac{1}{2}$ 71 77	62 90	55 70	62 71	56 72	17 71 $\frac{1}{2}$ 79 7 $\frac{1}{2}$	62 $\frac{1}{2}$ 70 7 $\frac{1}{2}$						
Group II-A DC													
1. Subject M	21 84 66 70	56 84	40 60	56 60	52 60	70 76	63 66						
2. Subject N	21 86 66 72	92 80	72 84	56 88	64 84	54 78	46 60						
3. Subject O	26 78 72 68	80 84	60 64	72 76	64 80	66 72	48 62						
4. Subject P	30 92 72 76	76 96	68 48	68 84	48 56	62 67	76 64						
\bar{X}	85 69 71 $\frac{1}{2}$ 76	86 60 64	63 77	57 70	10 $\frac{1}{4}$ 63 73 $\frac{1}{4}$ 10 $\frac{1}{4}$ 58 $\frac{1}{4}$ 63 4 $\frac{3}{4}$								

TABLE III

Individual and group mean scores for both active and delayed-comparison (DC) groups for the English Language Institute Placement Test, and both the pre-tests and post-tests in phoneme discrimination, mimicry, and oral reading. Column 1 indicates the age of each S.

	DISCRIMINATION TEST					PRODUCTION TESTS													
	I	II	III	IV		Mimicry	Reading												
Group II-B Active																			
1. Subject Q	26	66	66	56	64	60	56	68	44	52	68	58	64	50	54				
2. Subject R	20	56	50	54	44	92	60	76	80	80	60	68	60	78	53	60			
3. Subject S	23	70	68	68	80	92	56	72	84	88	52	76	66	76	64	76			
4. Subject T	36	70	58	58	80	80	60	60	88	88	64	64	64	74	58	62			
\bar{X}		65½	60½	59	67	81	58	66	80	75	57	69	7¼	62	73	11	56½	63	6¼
Group II-B DC																			
1. Subject U	31	82	70	74	68	68	56	48	52	56	44	60	60	66	58	64			
2. Subject V	33	52	48	60	68	72	56	60	72	84	56	80	56	70	51	66			
3. Subject W	25	82	66	72	76	84	48	64	84	80	60	68	64	72	65	70			
4. Subject X	20	80	70	80	80	80	68	76	72	60	68	72	74	82	60	74			
\bar{X}		74	63½	71½	73	76	57	62	70	70	57	70	5¼	63½	72½	9	58½	68½	10

systematic differences, although there is a trend, reversed in two instances, for the members of the three active groups to achieve superior results in all tasks. The mean scores for the active groups on the discrimination tests are greater than those of the matched delayed comparison group counterparts. Two of the active groups were superior in mimicry and two, but not the same two, were superior in the oral reading task. None of these trends is significant.

Since it was impossible to control for age given the limited number of Ss available for the study, it was decided to re-examine the results considering age as a possibly relevant factor. Again no systematic effects emerged. In short, no significant support for the hypothesis was discovered.

DISCUSSION AND SUMMARY

The students involved in this experiment were all highly motivated to learn English and had travelled to the United States from their native lands to do so. In this respect, they differ markedly from the high school students who have been the Ss investigated in the Keating Report, the Pennsylvania Project, and the Lorge-Bernard studies. All of the students followed the same academic schedule, were taught by the same teachers as their matched peers, and were exposed to identical amounts and durations of classroom and language laboratory materials. Attendance in the language laboratory was mandatory and absences were rare. All made satisfactory progress in improving their mastery of English, according to their instructors. In particular, pronunciation improvement as measured by the tests employed in this study was general, which indicates that the training they received in the classroom and laboratory was effective.

No significant support was found for the hypothesis that active students would make significantly greater progress in pronunciation discrimination and production than their matched counterparts who underwent a delayed-comparison cycle after completing each exercise of the pronunciation lesson. What trends did emerge clearly favored the active group in the predicted direction. It was not possible to control for age differences, but a re-examination of the data with this in mind fails to uncover any consistent effect. Only one of the Ss was female, so sex differences are irrelevant.

Only 24 Ss were involved in the study, and it is conceivable that the results of a more expanded study would be at variance with those of this experiment. The implications of the findings are momentous especially if the study is replicated. They imply

that when as many relevant variables are controlled as can be controlled, insofar as phoneme discrimination and authenticity of phoneme production are concerned, there is no significant difference in achievement between a delayed-comparison cycle or active exposure to pronunciation lessons for an equivalent amount of time.

REFERENCES

- Aleamoni, L. M., and R. E. Spencer. 1969. An evaluation of the Pennsylvania Foreign Language Project. *Modern Language Journal*, 53,421-8.
- Anderson, E. W. 1964. Review and criticism. *Modern Language Journal*, 48,197-206.
- Bernard, E. G., and S. W. Lorge. 1959-62. The use of the language laboratory in the teaching of French in secondary schools. Research supported by the New York State Education Department.
- Bernard, E. G., and S. W. Lorge. 1962-63. The relative effectiveness of four types of language laboratory experience. Research supported in part by the New York State Education Department.
- Bottiglia, W. F. (ed.) 1957. Report of the committee on the philosophy of the language laboratory. The Language Classroom Reports of the Working Committees, Northeast Conference on the Teaching of Foreign Languages. Cambridge, Mass.: M.I.T.
- Catford, J. C. (ed.) 1968. *Studies in Language and Language Behavior*, VI. Ann Arbor: Center for Research on Language and Language Behavior, The University of Michigan, Contract OEC-3-6-061784-0508.
- Clark, J. L. D. 1969. The Pennsylvania Project and the "audio-lingual vs. traditional" question. *Modern Language Journal*, 53,388-96.
- Company, E. 1968. Problèmes psychopédagogiques des laboratoires de langues: Première partie. In Catford 1968,279-400.
- Company, E. Electronic aids for teaching languages. Unpublished paper.
- Edling, J. V., et al. 1963. Reports on the Keating Report. American School and University, 36,33-8.
- Flaxman, S. L. (ed.) 1961. Modern language teaching in school and college. Reports of the Working Committees, Northeast Conference on the Teaching of Foreign Languages. New York: New York University.
- Gaarder, A. B., and J. C. Hutchinson. 1963. Brief analysis of the Keating Report. Washington, D.C.: U.S. Office of Education.
- Green, J. R. 1965. Language laboratory research: A critique. *Modern Language Journal*, 49,367-9.
- Grittner, F. 1964. The shortcomings of language laboratory findings in the IAR research bulletin. *Modern Language Journal*, 48,207-10.
- Hamilton, S. 1970. On the conversation class. *French Review*, 43,474-9.
- Hayes, A. S. 1963. Language laboratory facilities: Technical guide for the selection, purchase, use, and maintenance. *New Media for Instruction* 4. Washington, D.C.: U.S. Department of Health, Education, and Welfare.
- Hocking, E. 1964. Language laboratory and language learning. Department of Audiovisual Instruction, National Education Association of the United States, Monograph No. 2.
- Hocking, E. 1969. The laboratory in perspective: Teachers, strategies, outcomes. *Modern Language Journal*, 53,404-10.

- Keating, R. F. 1963. A study of the effectiveness of language laboratories. New York: The Institute of Administrative Research, Teachers College, Columbia University.
- Kieser, W. E. 1964. Hints on the teaching of modern languages. *Canadian Modern Language Review*, Summer, 31-5.
- Lado, R. 1957. Test of aural perception in English for Latin-American students. Ann Arbor: The University of Michigan.
- Lado, R. 1964. *Language teaching*. New York: McGraw-Hill.
- Locke, W. N. 1965. The future of language laboratories. *Modern Language Journal*, 49, 294-304.
- Locke, W. N. 1960. To record or not. *Modern Language Journal*, 44, 278-9.
- Lorge, S. W. 1963. Report of study on foreign-language laboratories in secondary schools. *Audio-Visual Communication Review*, 11, 70.
- Lorge, S. W. 1964. Language laboratory research studies in New York City high schools: A discussion of the program and the findings. *Modern Language Journal*, 48, 409-19.
- Lorge, S. W. 1965. Comments on language laboratory research: A critique. *Modern Language Journal*, 49, 369-70.
- Mathieu, G. 1960. A brief guide to sound labmanship. *Modern Language Journal*, 44, 123-6.
- Mathieu, G. 1961. Recommendations on the learning which should occur in the language laboratory and in the classroom. *Publications of the Language Laboratory, Series Preprints and Reprints*, 5, 58-95. Ann Arbor: College of Literature, Science and the Arts, The University of Michigan.
- Mathieu, G. 1963. Editorial comment on Keating's "A study of the effectiveness of language laboratories." *ML Abstracts*, 11, 2.
- Morton, F. R. 1961. The language laboratory as a teaching machine. *Publications of the Language Laboratory, Series Preprints and Reprints*, 1, 1-24. Ann Arbor: College of Literature, Science and the Arts, The University of Michigan.
- Mueller, T. 1958. Perception in foreign language learning. *Modern Language Journal*, 42, 167-71.
- Otto, F. 1969. The teacher in the Pennsylvania Project. *Modern Language Journal*, 53, 411-20.
- Porter, J. J., and S. F. Porter. 1964. A critique of the Keating Report. *Modern Language Journal*, 48, 195-7.
- Smith, P. D., and E. Berger. 1968. An assessment of three foreign language teaching strategies utilizing three language laboratory systems, Final Report, Project No. 5-0683, Grant No. OE-7-48-9013-272, U.S. Department of Health, Education, and Welfare, Office of Education.
- Stack, E. M. 1963. The Keating Report: The case of the missing subtitle. *The ETL Newsletter*, 4, 1-7.
- Stack, E. M. 1964. The Keating Report—a symposium. *Laboratories: The effectiveness controversy*. *Modern Language Journal*, 48, 189-94.
- Valette, R. M. 1969. The Pennsylvania Project, its conclusions and its implications. *Modern Language Journal*, 53, 396-404.
- Watkins, J. M. 1960. The library system and the language laboratory. *French Review*, 34, 60-6.