General guidelines for test preparation and use of the language laboratory are developed in this paper. Comparisons between written tests and oral language laboratory tests underscore the importance of clarity in test item preparation. This study examines: (1) test construction, (2) types of questions, (3) test administration, (4) test scoring, (5) student reference sheets, and (6) post-test reviews. (RL)
Testing in the Language Laboratory

by Evelyn Uhrhan Irving

This paper does not intend to present the pros and cons of testing in the language laboratory.* To my mind, this type of testing has value and I will here point out certain aspects of it. Your individual language laboratory may not lend itself to being used in the manners to be described; however, it is hoped that from the discussion each of you might be able to find some basic principles to adapt to your own situation. For the administration of a specific language laboratory examination, I refer you to my article in the NALLD Journal of March, 1969, entitled, "A Final Examination in the Language Laboratory."

Successful LL testing is predicated upon the following:

1). The instructor must be convinced that giving such a test is valid. 2). He must know what he wants to accomplish through the LL test. 3). The students must have previously been consistently exposed to using the LL and be familiar with all aspects of its use. Without these prerequisites, the test is sure to fail from the standpoint of student performance, grades, or both.

The instructor must follow all rules for preparing and administering a good examination: questions should be appropriate and clear and grading should be possible within a reasonable length of time to insure early return of the results to the students. Grading a lab test is somewhat more restrictive than the

* Paper delivered at the Mountain Interstate Foreign Language Conference, Gatlinburg, Tennessee, Oct. 10, 1970
classroom test and this must be taken into account.

One should never assume that a lab test is equivalent to a classroom test. It has some similarities to a timed test, but let us consider the student taking a lab test. Ordinarily, taped material is played to him and he is expected to respond immediately to it in some designated way. Have you ever noticed the behavior of a student preparing to write the ordinary classroom test? Many, although not all, take the paper, leaf through it for the number of sheets and types of questions, decide which one to answer first, and begin. He may even, as you well know, begin with the last question even though you were careful to place the simplest questions first. The written test allows for these individual differences which cannot be so easily allowed for in a lab test.

The secretiveness of both form and content of a lab test makes me suggest that the first responsibility of an instructor to his students is to eliminate some of this suspense. The student should be informed prior to, and again at the time of and during the test, of the general nature of the questions and of the individual question parts.

There are three distinct steps to be considered in preparing any LL exam: construction, administration, and grading. All these must be considered as intricately related, for I recall a colleague who came to me excited about the wonderful oral test he had just given his students in the LL; a week later he came in desperation: would I please give him some idea how he could grade it!

Let us now examine briefly each of the steps already mentioned.
TEST CONSTRUCTION

The lab test will follow all patterns of any well-constructed test. A note of warning: DO NOT TEST IN LAB ITEMS WHICH CAN BETTER BE TESTED IN THE CLASSROOM. The construction will consist of making up the questions, preparing tapes, student sheets, and scoring sheets. If the student sheet contains instructions which are also heard on tape, be sure that these instructions are clear— and identical. Remember that preferably a student cannot replay the tape; he hears it once unless the tape repeats it.

What types of questions are suitable and desirable for a lab test? To my mind, none of them should be drills; they should be practical application of the language. Two aspects of language can be effectively tested in the LL: comprehension and speaking. Writing can be tested too, but this involves comprehension as well.

Before making the tape, a complete script should be written. The script will include all numbers, directions (in English and/or the foreign language depending on previous procedure in class and laboratory), indication of persons to voice the materials, as well as all pauses. The question arises—who should voice the test tape: you as the instructor, another instructor, a native, non-native, male, female, etc.? Here the matter must be decided within the realm of possibility and previous experience of the student. If he is already familiar with several voices (which he will be since he has used the laboratory), it is not so necessary now that he be "exposed" to other voices. Certainly a native speaker is better than a non-native, but a good non-native
better than a poor native. Clarity, pitch, reading ability, even- ness of volume, etc. are all important aspects of a recording voice. The voice, too, should be a "happy voice", one that will sound al- live, joyous, and not like a funeral dirge. If adequate voices are not available, with a patchcord and two tape recorders, or the equivalent, one can record from commercial recordings the makings of an excellent test.

Lest the teacher who does not have unlimited assistance in the LL "turn me off" here and now, let me suggest that all of this is not as complicated as it sounds. Depending on the number of sections to which a test must be given (or even if it is to be given to only one class), it is relatively simple to make up several versions of the same type of question at a time. Then by using various combinations (three forms of each of three different question types gives a total of nine tests), the instructor has several tests to last for the period of time in which the same text and laboratory materials are in use. Properly constructed tests may even be used with more than one text.

Here are some suggestions for the master test tape. The first and last item on the tape should be a built-in state- ment or question which illicits the student's name. An initial reading assignment might be preceded by a statement in the foreign language, "My name is ... ." As a final question, the last of a series of questions in the foreign language would be, "What is your name?" This not only serves to identify the student (twice for safety) but also marks the beginning and end of each stu- dent recording.
Another consideration is that pauses be built into the master test tape. They should be designed pauses. While making a dictation tape, someone can write during the time being established for the pause, adding a reasonably longer time to take into account the level of student ability. A rule of thumb for the length of pause for answering questions is twice the length of time required of the instructor. A third consideration is that the student always be given a signal for responding, especially when answering questions.

Editorial comments are helpful in the test tape. These might include, "All right, are you ready? Let's begin." As you approach the end of a section you might say, "We're almost finished with question number 3. The second-last item is ... ." The correct answer might even be inserted from time to time after the student has recorded his response, thus serving to orient the student. All instructions should be built into the tape, including those for removing headsets and thus eliminating that confusing moment when the tape has stopped and the students look around, wondering what to do next! Whether all of this is in English or the foreign language depends on the language level of the students and what the previous procedure has been.
Types of Questions

What kind of questions are suitable for the LL test? Brooks, Lado, Stack, Valette, and many others have devoted entire chapters to this subject. I can here mention only some types and how they can be handled.

Written responses include multiple-choice answers for minimal pairs (e.g., ship vs. sheep; bit vs. bet, etc.) which are to be identified by the student as the same or different sound. There might also be multiple choice or essay answers for a comprehension question.

A dictation (admittedly a rather unrealistic form of communication—so are minimal pairs, even in sentences, for that matter), has its advantages and sufficient to my mind to warrant its use, for it tests comprehension, structure, and spelling ability. For non-phonetic languages a dictation is much more valuable than for a fairly phonetic one. An experiment showed that a dictation for a French class, if not used previously, is effective as a final examination. It is my opinion that when all students score highly on a dictation, it is time to adopt other types of questions. Although a dictation can be given in the classroom, I generally prefer the laboratory which gives direct communication between the tape and teacher, eliminating extraneous noises.

Oral responses include the following. A passage is to be read orally; the student may be allowed to study the material previous to, or as part of the test. This oral reading checks the student's ability to voice the language (including fluency, intonation, stress, juncture, etc.). Scoring of such a question can be precise and valuable to instructor and student.
Questions dealing with personal matters, everyday affairs, material studied in class (an excellent way to integrate class and lab), or the content of a paragraph assigned for listening or reading during the laboratory test may be answered orally by the student on tape. If he is to answer questions on a paragraph he has read, this paragraph may or may not be before him as he answers the questions. Questions, for listening only, might precede the playing of the paragraph, repeating them later for response.

A picture may be supplied the student which he is to use as a basis for recording a description of it, an incident which it recalls to him, or responses to questions concerning it.

Dialogues provide another type of oral response. Those learned in class might be handled with the tape representing one speaker and the student the other. It is suggested that these be variations on the dialogue rather than checking of rote memory of the dialogue. The ingenious teacher can create a test involving a dialogue possible to a telephone conversation; in actuality, a telephone booth with only the speaker present is the closest approximation of the laboratory booth to a real life situation.

This same teacher can set up a dialogue between two students in electronically-connected booths to carry out a telephone conversation, an appropriate topic or choice of topics having been assigned. *Concept Approach to Spanish*, 2nd ed. by Zenia Sacks da Silva contains some tape exercises where the student participates in the conversation.
ADMINISTERING THE TEST

As mentioned previously, all instructions for the test as well as all pauses should be part of the master test tape. In addition, all equipment should have been previously checked to ensure proper operation and each position made ready. It is wise never to plan to give the test to as many students as you have recording positions, since inevitably, one of them will fail to function. A seating chart is helpful to identify students as well as malfunction of any of the equipment.

With the students now physically present for the test, the atmosphere of ease attempted to be recorded into the tape should be brought into the LL. Sometimes appropriate music over a loudspeaker is beneficial. Since "to err is human," the student might be instructed that if he recognizes he has made an error during the recording, he may correct it within the allotted time for that response.

Testing is simplest for those who have a LL with recording at each position and all machines controlled from the console. Each tape can be programmed from the console and the student machines put into operation only during student response. This eliminates all need for the instructor to be hearing the instructions or questions when checking tapes and thus provides for minimum time in grading.

Student-controlled machines require some effort from the student. However, since students should be accustomed to working with the LL equipment before they are subjected to the test, they
will already be familiar with the equipment and can probably handle it with more ease than the instructor frequently does. Machines with pause buttons are very helpful in this case and anyone planning a language laboratory might keep this in mind. Of course, the procedure here is that the student put his machine into record operation, only when responding.

When testing according to the procedures outlined above, the instructor must decide whether all parts of the test are to follow immediately upon each other, or whether there should be a rest break between parts.

Even the most unsophisticated form of LL—one or two tape recorders—can serve to test students orally. Scheduling will be different from above and total testing time much longer. With this method, dictations require only a tape recorder and possibly headsets. Reading exercises such as those outlined above require basically that the student record the prepared material. Grading with this equipment is actually easier than grading a separate tape for each student since the student reading can be made consecutively on the same tape.

Question and answer tests can be done with two tape recorders, the student hearing the questions from one recorder and recording his responses on the other. Proper pauses can be built into the master tape and the student activate the recorder for response during the pause. As many students as tape recorders (minus one) can take the test at the same time.
Evaluation

A scale system of scoring makes oral grading essentially objective. Those who monitor their students in the LL or class, identifying the type of error, have undoubtedly already developed scales for this. There are examples of this type of scoring in Rebecca Valette's book, *Modern Language Testing*, as well as in many other. Once an instructor understands the principle, he can easily arrange his own scale for scoring.

Let us take as scoring example, the foreign language answer to a question in the foreign language. Let us suppose that we decide on a 3-point scale for each question. The scale might be:

<table>
<thead>
<tr>
<th>Points</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>perfect reply</td>
</tr>
<tr>
<td>2</td>
<td>good, but with faults</td>
</tr>
<tr>
<td>1</td>
<td>understood question; unacceptable reply</td>
</tr>
<tr>
<td>0</td>
<td>no reply</td>
</tr>
</tbody>
</table>

When a similar scheme was used, the scale for scoring having been previously established, it was found that teams of instructors rarely disagreed in the evaluation of oral replies.

A scoring scheme for a reading passage might be set up as follows: Break points into separate evaluation for 1). Pronunciation and 2). General reading ability (fluency, intonation, juncture, etc.); combine this to give a total grade for the passage. As an example, if a reading passage of fifty words is to count ten points, a suggested point distribution might be as follows:
Fluency, intonation, etc.: 4 points, distributed,

4 points  perfect performance
3         good but not perfect
2         fair but with definite faults
1         poor; total response, but unacceptable results
0         no response; partial, incomprehensible response

Errors in pronunciation (all errors within a word count as one): 6 points, distributed,

6 points  0 errors
5         1-3
4         4-6
3         7-9
2         10-12
1         13-15
0         16 or more

A reading passage of fifty words (or even less) appears to give a good evaluation.

If one were to use a scoring system such as this for an oral composition, the scoring can be done in much the same manner as for a written composition. A certain portion of the total value of the question can be assigned to:

A. Content
B. Structure, or grammar
C. Pronunciation, fluency, etc.
Student Reference Sheets

Sheets should be designed for the written portion of a LL test, and for indicating errors and correct answers to the oral portion.

Multiple-choice answers or the enumeration of oral choices (a, b, c, etc.) can be printed on a sheet. Space for a dictation can be allowed, numbering sentences for easier identification. The reading text can be part of this same scoring sheet and errors marked as they are noted by the scorer.

Another sheet can be made with the oral questions and appropriate full or partial answers. On this the scorer writes in the student error while checking the tape; when the sheet is returned to the student he can read his error as well as the correct answer. Dictations may also appear on this same sheet.

All the above, although requiring some time to prepare, will make scoring much easier and also provide visible evidence of the errors the student needs to correct.
POST TEST REVIEW

A laboratory test requires review in the same manner as a classroom test.

If assistants are available, they might monitor the tape with the student, pointing out corrections where necessary. Another procedure is for each student to listen to his own recording with the sheets before him indicating his errors. If he cannot correct his error, he can then ask help of the instructor or assistant. A good teacher is usually familiar enough with the student and the problems of the language to be able to identify the error by checking the scoring sheet.

Dictations can be handled as a class project in the classroom or laboratory, having available a printed copy of the dictation against which the student may check his answer while listening to the tape. Merely checking written forms is of little or no value.
CONCLUSION

This brief overview of testing in the LL leaves a great deal unsaid. However, it is hoped that it does lay out the general notions of this form of testing.

To my mind, although LL testing lacks the personal touch the instructor has with the student if he tests each one individually in a face-to-face situation, LL testing has definite advantages. Given sufficient equipment, it can be a great time-saver. If equipment is limited, it still saves the instructor's problem of scheduling each student individually and his time in examining him with all the details of directions, asking questions, etc. The taped test is uniform; each student receives it in the same way. Grading is objective: the teacher is not influenced by the bodily presence of the person and he preserves and can review the student response.

If one does not have opportunity to make a LL test with several types of questions, he can always begin with at least one and increase the number and variety of questions as he becomes more proficient. As teachers continue to use the language laboratory, systems of testing will be improved. This is an area which still needs great thought and development.

Evelyn Uhrhan Irving
Gareen-Newman College
Maryville, Tennessee

October, 1970
Selected Bibliography on the Language Laboratory


Selected Bibliography on the LL (page 2)

3M Company, *Taking the Mystery out of the Language Laboratory*, Pts. I, II, III, St. Paul, Minn. 3M Company


Journals and Reviews:

Audiovisual Instruction (publication of AECT), 1201 Sixteenth St., N.W., Washington, D.C.

Foreign Language Annals (publication of ACTFL), 62 Fifth Ave., New York, 10011.


NALLD Journal (Newsletter of the National Assn. of Language Laboratory Directors), ed. C.P. Richardson, Ohio University, Athens, Ohio 45701.

Many of the publications listed contain extensive bibliographies.