28. AUDITORY PRESENTATIONS AND LANGUAGE LABORATORIES

Steven D. Tripp
UNIVERSITY OF AIZU, JAPAN

Warren B. Roby
WASHINGTON STATE UNIVERSITY

28.1 INTRODUCTION AND HISTORICAL OVERVIEW

The use of mediated audio for instructional purposes is nearly a century old, and there is a deep history of studies of audio media in education, but in recent years the subject has been relatively dormant. Perhaps this is because technological advances in video and computing have led researchers in the direction of visual and textual presentations. Whatever the reason, much of the literature on this topic is not new. This chapter tends to emphasize work that cannot be called up-to-date.

This chapter is presented in two parts. The first part reviews research and makes recommendations about the design of auditory instructional materials. The second part of the chapter reviews research and makes recommendations about the most common application of auditory instruction, language laboratories.

28.1.1 Scope of the Auditory Instruction Review

This review of the literature on audio instruction is limited in its scope. It covers the use of audio, whether live or through conventional media such as tape, radio, or telephone. It does not address the use of audio for music instruction, nor does it include the applications of audio for various special education situations where students are suffering from handicaps that limit the effectiveness of other conventional media. Additionally, no attempt is made to include the use of nonverbal audio such as sound effects.

28.2 MEMORY FOR AUDITORY PRESENTATIONS

28.2.1 Memory Theory

A short description of current thinking may be helpful in order to put auditory memory in perspective (see also 3.2, 3.3). Potter (1990) distinguishes between three kinds of short-term memory: iconic, very short-term conceptual, and short-term verbal memory. Potter notes that these three types of memory are subject to differing types of interference and differing rates of forgetting. Iconic memory serves as a short-term visual buffer. It recognizes objects in 200 to 300 milliseconds and then holds that information for about 200 milliseconds. Very short-term conceptual memory receives scenes from iconic memory and stores them temporarily in conceptual form. These memories are not interfered with by iconic memory, but may be by other conceptual tasks. Very short-term conceptual memory decays in about 1 second. If these concepts are not embedded into a train of thought, they are lost. Finally, we have short-term verbal memory. This memory consists of an auditory buffer, which is severely limited in its capacity but may be maintained by continuous rehearsal. Its capacity is about 1.3 to 1.7 seconds, and its decay rate is nearly the same, allowing us to "replay" words or sounds heard a short time before. Meaningful background noise can interfere with verbal memory, and because this buffer is used for understanding written text, interference between the two channels may be experienced. This relates to the multichannel theories of learning that will be discussed later.

Potter (1990) stated that long-term memory may have three forms corresponding to the three forms of short-term memory. The dominant form of long-term encoding is conceptual, but visual and auditory long-term memories are also available. One implication of this conception of memory is that the extremely limited capacity of short-term memory implies that we should limit the number of chunks to be associated at any one time to three or four and that these *not* be redundant, because the short-term conceptual and verbal buffers are so vulnerable to interference. There is no reason to believe that information stored in any of one of the three short-term memories is more likely to be remembered because it was in that memory.

28.2.2 Early Comparison Studies

There have been many attempts to compare the relative instructional effectiveness of material presented through an audio or through a textual or visual medium. Erickson and King (1917) performed one of the earliest studies of the effectiveness of the auditory medium. Four groups of students from third to ninth grade were chosen, and each group was divided in half. One-half received the lesson from silent reading, whereas the other half was given similar material orally by the teacher. The following day, the order was reversed as to which half would read or listen. This procedure was followed two more times, for a total of four lessons. At the third- and fourth-grade level, the median score of the oral group was much higher than the medians of the group that read. At the fifth- and sixth-grade levels, the results were inconclusive. At seventh-, eighth-, and ninth-grade levels, the medians of the oral groups were much higher than in the groups that read the lesson. Needless to say, in these early studies experimental design was not what might be desired. Specific variables such as teachers' skill and interest were not taken into account, nor was the subject matter. Also, the only datum that was examined was the median score of the group.

Worchester (1925) tested 13 students to see whether oral or reading presentation of connected material was remembered better after 1 day, 2 days, and 7 days. In spite of methodological shortcomings, the author concluded that neither medium has a marked degree of superiority in the rate of learning of meaningful connected material.

The relative effectiveness of visual and auditory presentations of advertising material was studied by DeWick (1935). A group of 73 college students was divided into two groups. Three presentations of audio and visual advertisements were based on real advertising found in current magazines and on radio. One group heard the advertisement; the other looked at it in the booklet form. The auditory and visual advertisements were different. Time was held constant. After the experiment, the students were asked to write the names of products in order of presentation and then to describe the content. Next, the groups were switched and the same procedure followed. Five months later, all students were given the opportunity to recall the advertisements presented in both media. DeWick reported that: (a) The auditory presentation was superior in the recall of trade names after a delay of from 5 days to 5 months; (b) in immediate recall, both were similar; (c) recall was progressively weaker with visual presentations; (d) recall of ideas expressed in the advertisements was greater with the auditory method than with the visual.

Stanton (1934) conducted a follow-up study of DeWick, whose work was unpublished at the time. The auditory presentation was given without embellishments, such as music, and the visual material was presented without illustrations or other attention-getting devices. The same procedure was followed as in DeWick's study. Recall and recognition tests were given at three time intervals: 1 day, 7 days, and 21 days. The auditory method of presentation was superior. Recall was highest for auditory presentations at 7 days, and recognition was highest at 21 days.

The above studies addressed memory, but they did not investigate comprehension. Young (1936) tested comprehension with 2,000 intermediate students from Iowa and Texas. Four modes were used to present the material: (a) The teacher read aloud to students; (b) teacher read aloud and students read the selection silently; (c) students read selection once; and (d) students read for the time allotted the teacher to read orally. At the end of each presentation, a comprehension test was given; a delayed test was given 1 month later. Young reported that the students got very little information from an oral presentation and remembered even less after a month. However, the oral presentation was more effective than either of the silent readings, both immediately and after a month. He also found that the children who do poorly in comprehending through reading also do poorly in comprehending through hearing.

Larsen and Feder (1940) asked whether psychological abilities differentiate between the processes involved in reading and listening comprehension. After hearing and reading selected materials that varied in difficulty, 151 students were given both reading and listening comprehension tests. On these tests, there was a superiority of performance on reading comprehension over listening comprehension. However, this superiority depended on the aptitude of the student. Students with lower aptitudes tended to show equal results in listening and reading. The higher-ability groups showed a superiority in the reading comprehension tests. The authors concluded that comprehension is a central function operating independently of the mode of presentation. This study foreshadows modern results that emphasize the relative effectiveness of auditory presentations for students, who, for age or other reasons, do not read well.

In general, these studies, whatever their shortcomings, indicate that auditory presentations can be at least as effective as live or print presentations and are practical alternatives to conventional instruction. Under certain circumstances, especially when the materials are well prepared, auditory presentations may be superior.

28.3 AUDIO TECHNOLOGY

Early educational technology such as the radio, the telephone, the tape recorder, and the loudspeaker attracted the attention of researchers. Loder (1937) compared the retention of factual materials presented over a loudspeaker system and directly by a speaker. Two groups totaling 449 students were rotated in the experiment. A pretest, a test immediately after the lesson, a test 1 day later, and one 20 days later were given. One group saw the speaker, and the other group heard him from another room. The direct group performed better, but later tests showed that the means were not significantly different.

Cook and Nemzek (1939) compared the effectiveness of radio instruction and conventional teaching. Two intermediate

school groups were match-paired by age, sex, and intelligence for this experiment. Pretests and posttests were given. Their results showed a superiority for the radio-taught classes.

Rulon (1943a), using phonographic recordings, conducted an experiment to compare the amount of information gained by students who listened and those who studied the same material in printed form. Time was equalized. A total of 418 students listened to the recordings; 426 students studied the printed material. All students involved took a pretest, a posttest, and a test 1 week after the experiment was completed. Separate *t* tests were used to compare means of the pretest and posttest and also the delayed test. According to Rulon, the study of the printed material was superior to the method employing the recordings. However, a comparison in tests taken after a week showed little difference in methods employed. From this result, he concluded illogically that recordings make more of a lasting impression than printed materials.

Rulon (1943b) later conducted a similar experiment to compare the amount of information gained by students using phonographic recordings with the amount gained by students who studied a unit incorporating the same material presented in a textbook. This experiment probably is closer to an actual classroom situation, although, in this case, the textbook was prepared using the recordings as a primary source. Instructional methods using the textbook were not controlled. Testing procedures were similar to those above. The results, also using the *t* test, showed that phonographic recordings failed to show any superior effectiveness in teaching the "informational" aspects of the lesson.

In a third study on the effects of phonographic recordings, Rulon (1943c) investigated the motivational values of recordings. Using the same recordings and textbooks prepared for the earlier experiment, two groups of students were given access to supplementary reading materials after one group had heard the recordings and the other had read the material. Motivation was measured by which group used more supplementary reading materials. A total of 193 students used the recording, and 187 used the textbook presentation. Rulon's study showed no difference between the groups in terms of motivation to use supplementary reading material.

Kramer and Lewis (1951) investigated whether there was a difference in memory and comprehension between two groups of students in which one group sees and hears the speaker and another group only hears him. In the visual group were 128 students, and 120 were in the audio group. Both groups were in the same lecture room separated by a large, heavy curtain. Loudspeakers were used, and the lecture was given simultaneously to both groups. After the lecture, both groups took the same test. Both groups had been told that grades would not be counted. Kramer and Lewis reported that the mean of the visual group was higher than that of the audio group and that the visual group had a wider range. They concluded that the speaker's visible action somehow contributed to the ability to understand and remember the ideas in the lecture.

Leshan (1942) used a phonograph record to determine if a habit could be broken by suggestion while the subjects were asleep. At a summer camp, three small groups of young boys, all fingernail biters, were chosen for the study. The experimental group heard the suggestion to stop biting nails played 300 times a night over a period of a month and a half. The control groups received no auditory messages. Leshan reports that, at the end of the experiment, 40% of the experimental group had stopped nail biting, whereas none in the control groups had.

Similarly, a series of research studies with tape recorders attempted to determine whether or not students could be taught while they were sleeping. Fox and Roblin (1952) selected 30 people who had no knowledge of Chinese. Pretests and posttests were given. A tape recorder with a pillow microphone and an automatic timer switch was used for one-half hour during the night. Matched Chinese words with English words were given. Three groups were chosen: group 1 heard Chinese words and true English equivalents, group 2 heard Chinese words with false English equivalents, and group 3 heard music only. There were several weaknesses in the design, including the fact that the students were not observed at night. Fox and Roblin reported that the following morning the group that had the true translations learned the same list much more quickly than the other two groups, with the false translation group taking the longest. The authors conclude that learning can occur during sleep.

Gallacher and Stevens (1954) found that the use of tape recorders to teach spelling improved performance from 50 to 100%. The published description of this study was incomplete. However, as a result of this report, other studies were conducted on the effectiveness of tape recording used with spelling.

Gibson (1958, 1959, 1960) reported a 3-year experimental comparison of a tape teaching program with conventional instruction at an Omaha, Nebraska, junior high school. Two areas were chosen to study the effectiveness of tape recording: spelling and conversational Spanish. Oral and written tests were used. Results of findings included the assertion that tape instruction was superior to conventional instruction when the criterion was in the number of words correctly spelled. Both methods were similar with respect to the recognition of words misspelled. Spanish classes taught by a non-Spanish teacher using Spanish tapes, and classes taught by a Spanish teacher were similar in achievement scores. The following conclusions were made: (a) Tape recording is an effective method for teaching conversational Spanish to seventh-graders; (b) regular classroom teachers can effectively teach conversational Spanish by means of tape prepared by Spanish specialists; (c) students can learn to spell as effectively with a tape as with conventional classroom procedures; (d) with proper orientation, large groups can be taught spelling effectively; and (e) teaching with tapes produced no adverse effect on attitudes toward the subject.

Popham (1962) studied the effectiveness of tape-recorded lectures in teaching a college-level education class. Thirty-six students were divided into 18 matched pairs. Chi-square analysis revealed no significant difference between

assorted variables. One group was a conventional lecture discussion; the other group was taught by tape-recorded lectures with student-led discussions. This experiment continued over one semester. Pretests in achievement and a test to measure student opinion were given. Both tests were repeated at the completion of the course. Popham reports that both groups had increased performance on the achievement tests; there was no significant difference between them. There was no significant difference on reactions to the courses. However, the opinions of the tape-lecture sections were generally favorable towards the technique.

In a similar study, Menne, Klingenschmidt, and Nord (1969) provided taped lectures, tape recorders, and printed notes to 209 college students. Another 408 students attended regular lectures. Overall, there was no significant difference, but students in the lowest quartile showed an advantage in the tape condition. Also the dropout rate was lower with the students using tape.

As a result of the interest in taped instruction, the National Center for Audio Tapes (NCAT) was established in 1955 and was known popularly as the National Tape Repository (DeKieffer, 1973). The repository originated at Kent State University and by 1960 housed approximately 2,000 titles. The NCAT was moved to the University of Colorado in 1960 because of the facilities for high-speed duplication that had been established. A survey of tape duplication facilities by NCAT in 1965 determined that of 350 institutions responding, 223 had tape-duplication libraries. The majority were in institutions of higher learning. In 1973 the NCAT had over 14,000 titles and duplicated more than 17,000 tapes.

In addition to the tape recorder, there was interest in the telephone as an instructional medium. Cutler, McKeachie, and McNeil (1958) conducted a study concerning the relative effectiveness of teaching via the telephone. Two matched groups of 10 were selected. One group was taught elementary psychology in the traditional manner, the other by telephone alone. No text was used, but a list of suggested readings was furnished. The telephone group was connected to a system in which all participants could speak to each other. Gains in knowledge were found in both groups, and there was no significant difference in the gain between the two classes. Although there was evidence of a novelty effect, the method appeared practical. Rao (1977) summarized the limited research on telephone teaching and concluded, ". . . the research done on the effectiveness of teleteaching indicates that teleteaching is an economical and effective tool" (p. 483).

In conclusion, audiotechnology such as the tape recorder, has been generally shown to be as instructionally effective in conventional teaching. In addition, such technologies lend themselves to special situations. The telephone obviously allows distance education where students are unable to attend classes. In addition, it appears that the tape recorder allows learning while asleep. Although it is difficult to assess the validity of these sleep-learning studies, they point out that auditory media may have applications under conditions when conventional media are not feasible.

28.4 AUDIO-TUTORIALS

There are few cases of the large-scale, systematic application of audio technology to instruction. One interesting and extensive implementation of audio for instructional purposes is Postlethwaite's "audio-tutorial instruction" (ATI). This approach, which has been widely reported (Postlethwaite, 1970, 1972, 1978, 1980; Svoboda, 1978; Button, 1991), is more a complete instructional system than just an application of audio, but its long history and wide application make it an important source of information. Because of its well-elaborated structure, I will describe it in depth.

ATI began almost by accident in 1961 at Purdue University (Postlethwaite, Novak & Murray, 1972), when Postlethwaite was attempting to provide supplementary materials for weaker students in freshman botany. Simple lecture material was made available on a self-study basis through the audiovisual department. During the semester, these tapes evolved into programmed experiences that directed the students' attention to sections of the textbook, pictures, and diagrams, as well as live plants. Eventually experiments were added, and the entire week's study could be covered without attending any formal sessions. Student reaction was favorable, so in 1961–62, an experimental section of 36 students was taught entirely by audio-tutorial. Results on the conventional exam showed that the experimental group performed as well as the regular students. Interviews with the students led to the creation of a completely restructured course.

In designing the new course, Postlethwaite studiously avoided using words like lecture, recitation, and laboratory, which he felt connoted formality and passivity. The new course consisted of "independent study sessions," "general assembly sessions," "small assembly sessions," which included "integrated quiz sessions," and other small assembly sessions. Other additional activities were sometimes included.

The independent study session is the heart of Postlethwaite's system. Students would check into a special learning center with 36 booths equipped with tape recorders, movie projectors, and other materials. Armed with a list of behavioral objectives and the tape, the student would proceed through a series of activities. The tape was not a lecture. Rather it was a kind of programmed guide that would direct the student's attention to certain parts of the book or other materials. The tape acted as a kind of tutor, suggesting activities and pointing out important information. The student would perform experiments, read short segments of text, fill in diagrams or charts, view films, or examine specimens. The program could be interrupted at any time, and students could talk with other students or instructors if they were confused. Students could also omit activities if they felt they understood the content.

The general assembly sessions were held at the end of the week and were intended to support activities that could

not be performed at the individual session, such as exams, guest lecturers, and long films.

The integrated quiz sessions had an especially interesting structure. About eight students would be seated around a table covered with items drawn from that week's objectives. On the assumption that the best way to learn about something is to teach it, the students would be asked to discuss an item chosen at random. The student would first name the object, relate it to the behavioral objective, and then explain or demonstrate the object to the class. Once all students had completed this graded, oral quiz, a 20-point written quiz was given.

Another interesting feature was that although a small project involving the collection and analysis of data was compulsory for all students, any students desiring a grade of "A" were required to complete a second, original project.

It appears that Postlethwaite was acutely aware that an audiotape-mediated course might take on the appearance of impersonality. To avoid this, several measures were taken. The tapes were made with an informal, conversational quality. The instructors in the learning center were apparently told to maintain a pleasant personal manner. The senior instructor spent 3 hours per week in the small quiz sessions, meeting with about 48 students. Another 3 hours were spent informally visiting the learning center. He also held a weekly coffee hour to which all students were invited. Finally, he held an open house (for 600 students!) at his home once a semester. Postlethwaite considered this emphasis on personal contact and a well-structured sequence of learning events to be the essential ingredients of ATI.

As a result of his experience, Postlethwaite and his colleagues (Postlethwaite, Novak & Murray, 1972) were able to provide guidelines for the production of audio-tutorial lessons: First, they emphasized the importance of behavioral objectives. All course objectives should be listed first and test questions should be written concurrently, if possible. Second, you should list on cards all activities needed to reach the objectives. This should include the medium of presentation of the activity. Third, you assign the activities to the different kinds of sessions, as appropriate. Finally, you arrange the activities in a proper sequence, taking into account prerequisite knowledge. Once this is completed, the program tapes can be produced. Postlethwaite also provided guidelines for this step: First, assemble the required materials for an activity and make a demo tape. It is helpful to have an average student present as you do this, because the student's questions will trigger necessary elaborations and will give the tape a true tutorial feeling. Next, transcribe the tape and edit it critically. This eliminates redundancy and imprecision. Finally, make the final tape from the edited script, which may have emphasis or other helpful markings added.

In the production of the actual tape, Postlethwaite gave the following guidelines: First, with regard to voice, use a conversational tone, speaking clearly and cheerfully. Vary your tone frequently. Speak rapidly. Avoid "uh's" and other annoying speech habits. Second, concerning the content, aim for critical thinking. Do not lecture. Do not read the directions. Involve the students by directing their attention to actual things. Keep things simple but don't repeat, because the students can play back the tape if necessary. Proceed from the known to the unknown. Keep the lesson as short as possible, but be sure to clarify critical points. A variety of voices and sounds can increase realism. Finally, with regard to the mechanics, adjust the volume and tone to an appropriate level. Keep a constant distance from the microphone. To make corrections, find a natural pause in the conversation and begin there. Signal the student to stop the tape while performing an activity. A sound effect is better here than a spoken instruction that will have to be repeated many times.

Since a good part of design is avoiding errors, Postlethwaite and his colleagues listed some pitfalls associated with ATI. Since these pitfalls may be encountered with any mediated course, they are worth repeating. First, they list problems associated with the structure of ATI. One problem may be the lack of specific objectives. The sequencing of instructional events requires the clear specification of objectives. Also, ATI cannot be a lecture-on-tape. This is certain to be more boring than the original. Visuals that are too complex, lengthy, or irrelevant can detract from learning. Study materials that are unnecessary to the objectives are wasteful of time and money. Auxiliary experiences outside the learning center can be very valuable, but discussion and evaluative feedback are essential. Synchronization of mediated lessons with actual activities is difficult to predict and requires trial-and-error adjustment.

There are also psychological issues associated with ATI. First, the structure of the course is crucial to learning. Postlethwaite recommended Ausubel's subsuming structures, but augmented this where appropriate. Equally important, substructures of the course should also be paced and subsumed appropriately. This is very difficult to predict in advance, and almost always requires adjustment as a result of student feedback. Because older students have extensive vocabularies, direct experience is not always necessary, but when new concepts are being introduced, direct experience supports meaningful learning. Student feedback is essential to determining whether instruction is effective. On the other hand, feedback on small objectives may not be useful because it does not reflect the larger structure of the course, which is intended to clarify difficult concepts by repeated experiences. Finally, affective factors are easy to over- or underestimate. However, they cannot be ignored. Individualized audio-tutorial instruction capitalizes on the students' desire to be in control of learning, but it does not satisfy their social drives. For this reason, group work needs to be included.

In summary, in spite of the age of the Postlethwaite book (1972), it can still be recommended as a guide for the individualization of courses, especially where they are to be audiomediated.

Despite the widespread use of ATI, evaluation studies are not numerous. The common results reported (e.g.,

Postlethwaite, Novak & Murray, 1972) indicate that it is as good as conventional instruction with positive attitudinal ratings. However, there are a few reports of an advantage for ATI. For example, in Huppert and Lazarowitz (1990), 15 students used an ATI unit to learn cell biology. Standard classroom lab techniques were used with 65 others. Although the control group had significantly higher pretest scores, posttest scores were equal, indicating greater gains by the ATI group Also, higher-ability students tended to score higher with ATI. In another study, Carter and Cooney (1983) reported that undergraduates exposed to modified ATI in a statistics class showed a significant increase in performance, even under low-motivational conditions. Low motivation was indicated by the fact that students generally expressed the belief that the course was irrelevant to their career. Despite some methodological shortcomings, Carter and Cooney's effect size was 1.8 standard deviations, an astonishing gain under the circumstances. However, this gain was against a control group that received no instruction. Thus, these numbers reflect the effectiveness of ATI, not its superiority to conventional instruction.

The most comprehensive review of ATI is Kulik, Kulik, and Cohen (1979). This meta-analysis of 48 reports of ATI found a small but significant achievement effect for ATI over conventional instruction. However, ATI had little effect on course evaluations or withdrawal rates. Also, aptitude and achievement correlated highly, indicating that ATI does not have a leveling effect as might have been expected with such a self-directed, self-paced approach. Thus, the best available evidence seems to indicate that ATI does have a small positive effect on achievement, but given the fact that it does not seem to have other advantages over conventional instruction, it appears that its use can only be justified when special circumstances apply.

One kind of special circumstance is when well-trained content teachers are not available. An excellent example of this is elementary science education. Joseph Novak (of the previous Postlethwaite group) developed elementary science lessons using ATI. In an extensive longitudinal investigation of the effects of ATI, Novak and Musunda (1991) reported a 12-year study of science concept learning. Twenty-eight of their best science concept ATI lessons were provided to 191 first- and second-grade children. Each lesson required 15 to 25 minutes to complete. As with the Postlethwaite materials, students were directed by the tape to interact with materials and pictures. The lesson were sequenced according to Ausubel's ideas about subsumption. Lessons were replaced after all students in the class had a chance to interact at least once. With few exceptions, all students were capable of proceeding through the lessons. Grade-1 students completed lessons 1 to 16, and grade-2 students completed lessons 16 to 28. Lesson 16 was repeated. All lessons covered basic science concepts and principles. These children were compared by means of interviews to 48 uninstructed students in grades 2, 7, 10, and 12. Early on, the interviews included the use of tests with pictures or diagrams. However, these were later abandoned because they were judged to be confusing or misleading. Data collection evolved into Piagetian clinical interviews that were then translated into concept maps. Concept maps were then graded for valid and invalid notions. Instructed subjects showed significantly more valid concept understandings and fewer misconceptions. A significant interaction showed that instructed subjects, over the 12 years, had a greater tendency to increase their number of correct concepts and decrease the number of incorrect concepts. This exceptional study strongly supports the validity of ATI even under conditions where the instructors are not well trained. Valid concepts were learned from ATI and evidently "scaffolded" more learning throughout the children's 12 years of schooling.

In summary, audio-tutorial instruction is the most complete and most well-documented method of auditory presentation. It has a general record of success, and, although it is not noticibly superior to conventional approaches, it has many valuable ideas, even for other nonaudio forms of mediated instruction.

28.5 COMPRESSED SPEECH

Because people were found to learn as effectively from tape recordings as from live speech, a variation on research on auditory comprehension emerged. This research was based on a simple observation: People can comprehend speech faster than narrators can speak. If a technology for speeding up the delivery rate of a tape recording could be developed, an apparent gain in instructional efficiency might be realized. This simple idea has generated a large body of research.

28.5.1 Processing Capacity

One way of measuring cognitive processing capacity is the comprehension of words per minute (wpm). Conversation typically takes place at a rate of 12 to 150 wpm (Benz, 1971; Nichols & Stevens, 1957). Since in conversation, one is simultaneously listening and composing speech, it was assumed that perhaps another 125 to 150 wpm of unused capacity might be available in simple listening. Because the rate for speed reading (Taylor, 1965) is 250 to 300 wpm, it was assumed that that much capacity might be available for listening. In fact Fairbanks, Guttman, and Miron (1957) found that rates of up to 300 wpm were possible with compressed speech.

28.5.2 Compression Technology

In early studies, compressed speech was produced by playing back a recording at a speed faster than that used to make the original recording. While this method is simple to produce, the vocal pitch and intelligibility were affected. The limitations of the "speed changing" technique generally rendered research findings questionable.

Miller and Lichlinder (1950) first demonstrated the tape-sampling method accomplished by deleting segments of the speech signal. A switching device was used that turned off the signal periodically. Garvey (1953) performed further experimentation in compressed speech by editing out segments of the audiotape and splicing the ends of the retained tape together. While Garvey's technique was successful, it was deemed too tedious except for research purposes. Fairbanks, Everitt, and Jaeger (1954) produced the first electromechanical apparatus that allowed both the expansion and compression of recorded tape.

Technological developments in the 1960s improved the Fairbanks technique. Scott (1965) utilized a computer to dispose of empty time intervals between words and to sample the time interval occupied by words differentially. While the computer proved to be the best means of producing compressed speech, the cost of computer time negated the application of this approach for other than experimental purposes. Electronic developments of the last 2 decades have now allowed the mass marketing of completely electronic compressor/expander tape recorders.

28.5.3 Comprehension

Numerous researchers have varied the rate of compression and measured the resulting effect on comprehension. Fairbanks, Guttman, and Miron (1957) found little difference in comprehension of selections compressed to 141, 201, and 282 wpm. Diehl, White, and Burke (1959) determined that listening comprehension was unaffected by changes between 126 and 272 wpm. Foulke (1962) used both literary and technical presentations and found that listening comprehension was slightly higher in the 175- to 272-wpm range than the 272- to 375-wpm range, at which point an accelerated loss in listening comprehension occurred. Foulke and Sticht (1967) measured a 6% loss in comprehension between 225 and 325 wpm, and a loss of 14% between 325 and 425 wpm. These and other subsequent studies (Boyle, 1969; Carver, 1973; Foulke, 1968; Foulke & Stich, 1969; Rossiter, 1970; Sticht, 1968; Wasserman & Tedford, 1973; Williams, Moore & Sewell, 1983–84) indicated that as word rate is increased beyond about 250 to 300 wpm, there is a decline in comprehension. Recall that Potter's verbal short-term memory buffer has a capacity of about 1.5 seconds. At 300 wpm that is about 7.5 words. Since these words must be processed as concepts to be understood, and our conceptual short-term memory has a capacity of only five to seven items, a limit of about 300 wpm seems reasonable. Indeed, Carver (1982) found that the optimal rate (which sacrifices some comprehension) for both reading and listening is about 300 wpm, indicating that there is an innate bottleneck in human information processing beyond which improved technology cannot take us.

However, numerous intervening variables must be considered before a determination of the optimum degree of compression can be made (Duker, 1974). Researchers believe that the ability of subjects to comprehend compressed speech may be dependent on the difficulty of the material. Readability has sometimes been found to influence normal audio comprehension (May & Lumsdaine, 1958; Chall & Dial, 1948), but others (e.g., Molstad, 1955) were unable to replicate this finding. Foulke (1962) determined that the comprehension of a scientific selection was less than the comprehension of a literary work at normal speed. However, at various levels of compression, the comprehension scores of the scientific selection declined less than those for the literary selection. This phenomenon may be because the comprehension scores for the scientific selection were lower at the normal rate; therefore the range in which they could vary was relatively small (Duker, 1974). Fairbanks, Guttman, and Miron (1957) investigated the effect of time compression on messages of various difficulty levels. The study seemed to indicate that, within the range explored, listening comprehension did not depend on the difficulty of the listening material. Goldhaber (1967) and Reid (1968) found that comprehension decreased as wpm increased, and that simplified material was better comprehended than more difficult material. Carver (1982) presented data that relate reading difficulty with efficiency. His data showed that for college students, listening to compressed materials written at the eighth-grade level produced the greatest efficiency. Raising the difficulty level of the materials caused comprehension to drop off abruptly.

Comprehension may also be measured in terms of delayed recall. George (1970) studied various rates of compression and two levels of material difficulty. He determined that more was forgotten at the lowest levels of compression in a delayed measurement, 1 day and 1 week following treatment. George indicated that although simplified materials were accompanied by some initial forgetting, the amount of forgetting with the passage of time was less than with more difficult materials. Friedman, Freedle, Norris, and Orr (1966) conducted retention studies among college freshmen and sophomores at speeds ranging from 175 to 475 wpm. After a 30-day delay, the posttest was administered again. The retention for the two highest rates, 425 and 475 wpm, were 117% and 90%, respectively, of their first-session scores. These authors reported that information presented in compressed format is retained as well as information that has not been compressed. Foulke (1966) pointed out that forgetting is not limited to the recall of information presented at accelerated word rates. In other words, there is no indication that compression, within broad limits, has a unique effect on the retention of information.

Length of presentation may also be a factor in comprehension and memory. Adelson (1975) examined comprehension by a group of college students listening to a 1-hour lecture at 175 wpm, as compared to the same group of college students listening to an equated 1-hour lecture compressed at 275 wpm for 40 minutes. Compressed materials produced less comprehension than did the normal rate materials. The author concluded that with compressed materials, the length of presentation appeared to be a critical factor, perhaps because of attentional fatigue or other factors.

Narrators of both sexes differ in vocal pitch, average word rate, variation in word rate, pitch, and loudness (Foulke & Sticht, 1969). Foulke (1968) examined the extent to which these factors interacted with word rate in determining listening comprehension. Three versions of a selection were presented to college students at normal and compressed rates by narrators of both sexes. Significant differences in scores on a listening comprehension test were associated with the word rate variable, but narrator's style did not interact with rate. Rossiter (1972) had college undergraduates listen to short informative messages by both male and female presenters. Students were tested on the content of the messages. Data analysis showed an interaction of the sex of the speaker with the sex of the listener, but the author dismissed its importance, concluding that the sex of the speaker was not of much consequence in determining listening scores of subjects participating in the study.

Learner characteristics may influence the comprehension of compressed speech. Some of these variables include the subjects' sex, age, intelligence, and reading ability. Duker (1974) determined that the comprehension scores of male and female subjects revealed no sex-related differences for word rates varying from 174 to 475 wpm. This conclusion is supported by other research studies conducted by Foulke and Sticht (1967), Orr and Friedman (1964), Ross (1964), Bell (1969), Ludrick (1974), and Klavon (1975).

Fergen (1954) and Wood (1966) found that ability to comprehend compressed speech increases with age and grade level of school children. However, beyond age 12, little difference is noted until age 60 or so. Goldhaber (1970) found significantly better comprehension for junior high school students over college freshmen and sophomores. Goldhaber attributed the difference to the interest level and the level of motivation for each population. Duker (1974) agreed that the effects of age and education on the comprehension of compressed speech cannot be generalized. Lysaght (1969) tested elderly, middle-aged, and young adult subjects. He determined that elderly subjects performed lower on a posttest measuring comprehension. Duker (1974) noted that the decline in the ability to comprehend compressed speech may result from ". . . changes in the central nervous system" as opposed to the variable of age, per se (p. 494).

Aptitude or intelligence may also interact with comprehension. Eckhardt (1970) used a 1-hour multimedia presentation at various rates of compression with Air Force recruits of varying aptitudes. Eckhardt concluded that test differences between the groups were due to aptitude and an aptitude-rate interaction. There was a comprehension loss for lower aptitude subjects at the higher compression levels. Sticht and Glasnap (1972) determined that low-aptitude men learned easier material better than more difficult material as a function of decreased wpm. High-aptitude men tended to learn material best at 175 wpm, independent of difficulty level. However, other researchers (Sticht, 1968; Watts, 1971; Williams, Moore & Sewell, 1983–84) found that subjects with lower aptitudes or lower reading ability performed as well at higher rates of compression as at normal rates. Fergen (1954) found no relationship between the IQ of grade school children and their ability to comprehend compressed speech selections. On the other hand, Goldstein (1940) and Nelson (1948) found a positive relationship between intelligence and comprehension of compressed speech. Cicardo (1974) examined the retention of a compressed speech message presented to junior high school students of various intelligence levels. Cicardo determined that IQ level affects factual retention of material presented, but there appeared to be no interaction with rate. Foulke and Sticht (1969) pointed out that the relationship between lower intelligence and the decline in comprehension may be attributed to the lower scores of less intelligent subjects, which have a lower variance. For this reason, Foulke and Sticht (1969) and Duker (1974) argued that the difference in comprehension cannot be attributed directly to intelligence level.

Reading ability may influence comprehension of compressed speech. Breed (1977) tested adult vocational technical school students to determine the differences in listening comprehension when subjects were categorized according to reading ability. The subjects in Breed's study listened to tapes that were time expanded and time compressed, varying in rate from 60 to 240 wpm. Breed indicated that listening comprehension and reading ability appear to be related to verbal skills. The poorest readers exhibited the poorest listening comprehension, and better readers were better listeners as measured by scores on tests of listening comprehension. Goldstein (1940) and Orr, Friedman, and Williams (1965) found a positive correlation between reading rate and ability to comprehend compressed speech. Conversely, both studies further determined that practice in listening to compressed speech resulted in an improved reading rate. Robertson (1977) determined that the comprehension of subjects is not affected when they are presented recorded materials within two reading levels below or three reading levels above their particular grade level. In general, it appears that a relationship between better reading ability and the comprehension of compressed speech can be established, although this may reflect an underlying verbal ability.

28.5.4 Training

It has often been speculated that practice might influence comprehension of compressed speech. This research has attempted to provide appropriate training prior to treatment in order to improve comprehension. Voor and Miller (1965) exposed subjects to five listening sessions at 380 wpm. Test scores indicated that comprehension increased as a function of exposure up to 7 minutes, and remained constant thereafter. Orr, Friedman, and Williams (1965) exposed blind subjects to listening material presented initially at 325 wpm and increased in 25 wpm intervals to a rate of 475 wpm. Subjects were tested for comprehension at 475 wpm and compared to equivalent pretraining test scores. An improvement of 29.3%

was noted. Friedman, Orr, Freedle, and Norris (1966) compared the comprehension scores of subjects given 35 hours of massed practice with test scores of subjects given 14 to 21 hours of distributed practice in listening to compressed speech. The authors concluded that the comprehension of the distributed-practice group was as good or better than the comprehension demonstrated by the mass-practiced group. Duker (1974) suggested that gradually increasing the wpm rate might have some benefit on comprehension of compressed speech. Klavon (1975) tested this idea, without effect, in at attempt to provide a controlled transition period. In general, studies (Foulke, Amster, Bixler & Nolan, 1962; Friedman, Orr & Norris, 1966) have found that although no particular method of training or practice appears to be any more effective than another, even small exposure to compressed speech can improve comprehension.

28.5.5 Affective Factors

There are some interesting affective factors to consider when using compressed speech (see also 34.7). Listener attitudes toward the speaker are improved significantly (Maclachlan, 1982). Maclachlan notes that people associate fast, fluent speech with confidence, knowledge, and enthusiasm. Because attitude learning is influenced strongly by feelings toward the speaker, compressed speech may have an unexpected application in such situations. Additionally, college students prefer listening to compressed speech over normal tapes (Short, 1977), apparently because of the time savings. Also, for students under about 14 years old, there is a preference for listening over reading, presumably because of their slow reading rates (Boyle, 1969).

28.5.6 Efficiency

Recall that the original impetus for speech compression was potential efficiency. The instructional implications of using compressed speech for efficiency are limited. When the time saved in compression was used (Fairbanks, Guttman & Miron, 1957) to elaborate certain parts of the text, comprehension for that part of the text increased. But as Sticht (1971) pointed out, the time saved in compression was lost in elaboration, and overall comprehension was not improved. Similarly, hearing the same text twice at double speed resulted in no more learning than hearing the same text once at normal speed (Schramm, 1972). Thus it appears that the instructional use of compression cannot be based on efficiency arguments.

28.5.7 Summary

In conclusion, a great body of research has been done on speech compression. It has confirmed the hypothesis that listeners can process speech at a much higher rate than normal conversational speech, with some loss of comprehension. Usually, any short exposure to compressed speech will result in improved comprehension. In general, no differences were detected for sex, or age between about 12 and 60. However, some differences in the ability to comprehend compressed speech may be due to aptitude or verbal ability. Compressed speech may be preferred to normal speech and may cause positive attributions to the speaker. However, the early hopes that it could lead to more efficient instruction appear to have been unjustified.

28.6 DISTANCE AUDIO EDUCATION

28.6.1 Radio

Although the history of the use of radio in education (see 13.4.2.8) is long, there is not a plethora of empirical data concerning its effectiveness (see also 13.4). Saettler (1990) gives an institutional history of early educational radio in the U.S., but does not mention more than a few empirical studies. After World War II, interest in instructional radio declined, so the situation has not improved.

Educational radio began at the University of Iowa in 1911 (Wolcott, 1993). The University of Wisconsin followed in 1919, and the Ohio School of the Air was established in 1929. Other "schools of the air" were established at many institutions, but the Wisconsin effort appears to have been the most successful, surviving until this day. Various attempts were made to evaluate the effectiveness of these broadcasts, and Saettler (1990) reports on two important ones. The first, the Ohio Evaluation of School Broadcasts Project, enumerated objectives that might be met by radio. Armed with these materials, the CBS American School of the Air was evaluated in 1940–41. It was determined that teachers who used the broadcasts found the broadcast extremely valuable. In spite of this, various defects were identified. First, purposes were not always well defined. Second, there was a need for printed aids. Third, there were errors of content selection which reflected a lack of understanding of the backgrounds of the pupils. Fourth, although the broadcasts were enjoyable, they were not as interesting as they might have been.

The second evaluation study, the Wisconsin Research Project, among other efforts, compared radio instruction with conventional instruction in six subject areas from 1937 to 1939. The differences were not significant, and only with music was the small difference in the direction of radio. Although there are not many other studies, those that exist (e.g., Constantine, 1964; NHK, 1956; Cook, 1964) typically showed that radio students performed at least equally as well as live audiences.

Bates (1983) reported that the British Open University's experience with radio instruction showed that the broadcasts tended to help the weaker students more than the successful ones. It must be remembered that Open University courses were taught primarily through correspondence texts and that the radio broadcasts were intended to supplement the texts. It is not surprising that students who found the texts difficult would welcome the added help afforded by the radio.

Radio is no longer widely used in education in the wealthier countries, but Wolcott (1993) reports that interactive radio is still used where long distances are involved, as in Alaska and Australia. In the poorer nations, radio is used widely because it is cost effective compared to other more sophisticated technologies. Radio is not expensive to produce or receive and can cover long distances using the AM band. The Agency for International Development (no date) reports that from 1974 to 1990 radio instruction programs existed in Nicaragua, Thailand, Kenya, Nepal, Dominican Republic, Papua New Guinea, Honduras, Bolivia, Lesotho, Costa Rica, Ecuador, Belize, Swaziland, and Guatemala. Reported results are encouraging. For example, in Bolivia effect sizes as large as .91 were reported for radio math as compared to traditional math. By dividing effect size by cost per pupil, a measure of cost effectiveness was obtained. According to data from the Agency for International Development, interactive radio was generally much more cost effective than textbooks or teacher training. In a related case study of radio-assisted community basic education (Eshgh, Hoxeng, Provenzano & Casals, 1988) in the Dominican Republic, gains in both math and reading as a result of radio instruction are reported.

In summary, as with other media, radio instruction has been found to be at least as effective as conventional instruction, although the literature is limited. Under certain circumstances, such as when conventional instruction is inadequate, the radio can be a cause of improved learning. Given the cost and reach of radio, it appears to be a viable medium in places where other media are too expensive or unavailable.

28.6.2 Other Media

In distance education, other recent audio-based technologies, besides radio, are now entering the field. Audioconferencing, both audio only and audioconferencing with images or data, is now possible through existing public telephone lines. Audiographic technologies such as the fax machine, the electronic blackboard, and both still and motion video images are becoming common. Use of high-speed data compression techniques will make these media more and more convenient and cost effective. Although these technologies are audio based, in the sense that they use telephone lines, they are not strictly speaking auditory and so will not be discussed here. A very complete study of these technologies and their limitations, along with suggestions to overcome those limitations, has recently been published (Wolcott, 1993), and the reader is referred to that source for further details (see also 13.4).

28.7 AUDITORY LEARNING VS. PRINT PRESENTATIONS

Since the general topic of multichannel presentations is covered in another chapter (see 29.3), it will not be discussed here. However there is some value to examining the specific question of the effects of learning the same material by listening or by reading. Olsen and Bruner (1974) stated that, "Each form of experience, including the various symbol systems tied to the media, produces a unique pattern of skill for dealing with or thinking about the world" (p. 149). Similarly, Salomon (1979) stated that, "Media's ways of structuring and presenting information—that is, their symbols systems—are media's most important attributes when learning and cognition are considered and should serve as the focus of our inquiry" (p. 216). These seem to imply that there should be some cognitive difference between listening to or reading a text.

This notion seems to contradict Clark's well-known hypothesis that media do not influence achievement. After an extensive review of the literature, Clark (1983) concluded that, "The best current evidence is that media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition" (p. 445). He has defended his position by arguing that even where apparent differences are detected, they are rendered dubious by various forms of confounding. Some of the forms of confounding that Clark cited are the novelty effect, the "John Henry Effect," unequal instructional strategies, unequal opportunity to learn, and unequal quality of instructional design.

A comparison of learning from print or audio seems to provide a convenient test of Clark's hypothesis, because it is relatively easy to control both content and instructional strategy. Since both audio and print are used to convey verbal information, the essential difference between a print or audio presentation (if otherwise equal in quality) is precisely the medium of delivery.

28.7.1 Audio and Print

Reading and listening seem to demand the same underlying linguistic competence (Mosenthal, 1976–77). In general, comparisons of learning from audio and print have shown no difference, as Clark would predict. Nugent (1982) and Rohwer and Harris (1975) found no differences for children, while Nasser and McEwen (1976) found no difference for college students. Needless to say, this assumes that the printed text is within the reading ability of the student.

Other research on reading and listening sometimes suggests an advantage for dual-channel presentation. Some studies showed that recall was higher when the very same material was read and heard than when it was presented in one channel alone (Hartman, 1961; Nasser & McEwen, 1976). However, Nugent (1982) found no advantage with children for print plus audio over either print or audio alone.

Recent research in England may shed light on this question. The research has largely been done by a pair of scholars, Furnham and Gunter, and their colleagues while studying the effects of mass communication on the recall of factual information. These researchers are apparently unaware of Clark's hypothesis and thus have not addressed it.

Although Furnham and Gunter have found a variety of results under differing conditions, they have consistently

found that subjects remember material presented in the print medium better than identical material presented in the audio medium or the audiovisual medium. Gunter (1987) concluded that this was due to the inherent capacities of these different media to convey knowledge. These differences were found for samples drawn from populations of schoolchildren, university students, military personnel, and nonstudents. Across these categories, the most consistent result was that subjects remember better from print materials than audio or audiovisual materials. Table 28-1 (adapted from Furnham, Gunter & Green, 1988) illustrates their findings.

Furnham and Gunter's research methodology may be best exemplified by their 1986 article (Gunter, Furnham & Leese, 1986). An approximately 5-minute television broadcast was delivered to university students in three groups. Group 1 received the TV broadcast in its normal form. (We are not concerned with this treatment here, but it is included for completeness.) Group 2 received the broadcast with audio only. Group 3 received a transcribed script of the broadcast. Time was held constant in all treatments. Memory was tested with free recall, cued recall, and multiple-choice questions, in that order. Results, as reported above, indicated that retention was always superior with print materials. Note that the content was designed for TV and that the print condition was the least "natural." These methods and results are typical. It is important to note that because these methods appear to be more rigorous than many of the earlier studies, and because the results have been replicated many times, they should be given greater weight.

Although Furnham and Gunter repeatedly found significant differences in achievement with print and audio, their work may have confounding factors. It is important to examine it in terms of Clark's list of possible confounds. First, it seems that the "John Henry" effect, the unequal instructional strategies, and the unequal quality of design arguments can be dismissed as irrelevant. No teacher was involved, so the "John Henry" effect cannot have occurred. Second, the content delivered was identical in terms of text and time; therefore no differing strategy or quality could be involved. Again, remember that in many cases the print version was derived from an audioscript, so, if anything, the print version should have been "inferior" in quality.

Different media may not be equally "novel" or familiar. But in this case, it is unlikely that either print or audio would be perceived as novel, so this objection seems weak.

Possibly, reading may be superior to listening because of greater opportunity to learn. Although Furnham and Gunter held total time constant, it is normal that people can read faster than an announcer speaks. Thus the subjects in the print condition may have been able to read the text more than once. Clark mentions this "reviewability" problem in his original article, but this objection seems to be unwarranted. If print is inherently "reviewable," then it is precisely this quality of the medium itself which influences student achievement. Thus media do influence achievement.

In a recent study, Tripp (1994) tested the differences between audio and print in a direct comparison that attempted to hold other factors constant, including reviewability and novelty, by presenting the same text by computer through either the audio or video (printed text) medium while holding speed and reviewability. As with Furnham and Gunter, the students who read the text remembered significantly more correct semantic units than the students who heard the passage. Clark (personal communication) has suggested that affective factors such as those mentioned by Salomon (1984) may account for the differences, but my students expressed the belief that learning from "tapes" was more difficult than learning from "books." Contrary to Salomon, they also expressed the belief that they would

TABLE 28-1. AUDIO AND PRINT (Adapted from Furnham, Gunter & Green, 1988)

Content	Source	Ss	Results*
News	Gunter, Furnham & Gietson, 1984a	128 school children	P > A = AV
	Furnham & Gunter, 1985	68 students	P > A > AV
	Gunter & Furnham, 1986	117 military students	P > A = AV
	Furnham & Gunter, 1987	101 adults	P > A > AV
Political broadcast	Gunter, Furnham & Leese, 1986	65 students	P > AV > A
Ad	Furnham, Benson & Gunter, 1987	69 students	P > AV > A
Magazine program	Furnham, Proctor & Gunter, 1988	63 students	P > AV > A
Science program	Furnham, Gunter & Green, 1988	60 students	P > AV > A (free recall) P > AV = A (cued recall)
Insect biology text (easy and hard)	Furnham, Gunter & Green, 1988	60 students	P > AV = A (free recall, hard text) P > A > AV (cued recall, hard text) P > AV > A (free recall, easy text) P = AV > A (cued recall, easy text)

*P = paper; A = audio; AV = audiovisual.

also expend more effort with books. Thus, if these attributions transfer to audio and text presented by computer, students may have tried harder with text, but there was no external indication of this. In any case, whether achievement differences are attributable to intrinsic qualities of the medium or student attributions of intrinsic qualities, or an interaction between the two, in this highly controlled case student achievement did vary as a function of the medium of delivery to the detriment of the audio presentation.

Given the robust results demonstrated by Furnham and Gunter and replicated by Tripp with differing content and audiences, it seems that it must be concluded that print is superior to audio as a presentational medium, when the content to be delivered is to be held constant.

28.7.2 Conclusions about Auditory Instruction

There is a considerable body of evidence that audio presentations, through a variety of technologies, can be at least as effective as other forms of instruction, and audio-based media may be usable when other media or live presentations are impossible or impractical. However, the Furnham and Gunter studies indicate that when audio and print are identical and time is held constant, people learn more from print. Under normal instructional circumstances, however, this is not the case. Typically, audio presentations are adapted to suit the instructional situation. Under these modified conditions, there is no reason to doubt the effectiveness of audio. The best example of this is Postlethwaite's audio-tutorial method which uses audiotapes as a kind of programmed guide in a larger instructional system. This system has been widely implemented and may serve as a model for other forms of individualized instruction. Finally, although there has been a great deal of interest in compressed speech for instructional purposes and it has been demonstrated to be effective, its practical applications are limited.

28.8 LANGUAGE LABORATORIES

28.8.1 History

Because of the nature of their task, foreign-language educators have been heavily involved in the use of audio equipment. They welcomed the first audio device, the phonograph, and have immediately adopted other advances in audio technology such as magnetic tape and digital media. Unfortunately, the history of the use of audio technology to teach languages has not been duly noted by historians of educational technology. Paul Saettler, in his definitive *The Evolution of American Educational Technology,* only makes passing references to foreign-language teaching, and language laboratories are granted merely one paragraph (p. 187). It will be demonstrated that this disregard is startling in view of the extensive use of, and massive investment in, audio equipment by foreign-language educators. Moreover, it will be shown that the research that accompanied these commitments has not been appreciated by the larger educational technology community.

28.8.1.1. History of the Use of Audio Resources in Foreign-Language Education: 1877 to 1945. Leon (1962) and Peterson (1974) have documented the early use of audio recordings by foreign-language educators since the invention of the phonograph by Thomas Edison in 1877. By 1893 there were commercial record sets available for Spanish and English as a foreign language. The phonograph was used in regular classes and for self-study at home, but to what extent is difficult to ascertain. In their 340-page annotated bibliography of "modern" language methodology (the references commence in 1880s), Buchanan and MacPhee (1928) include only nine entries concerning the phonograph. Three of these are listings of recorded courses; none of the six articles is a controlled study of the merit of the phonograph. The 491-page Bagley et al. volume (1930) contains no mention of the phonograph. This paucity of references is surprising when one considers that in the 1880s the field of phonetics was born out of the effort to teach proper foreign-language pronunciation. The literature of the period is full of articles on phonetics, and many pronunciation textbooks and teaching materials were published. One would have expected greater enthusiasm in the language-teaching community for the equipment that could provide native speaker models.

According to a contemporary (Keating, 1936), initial use of the phonograph and other devices such as the stereopticon (an early slide projector) was haphazard, and interest waned because there was no real absorption of modern inventions into the teaching program (p. 678). The Depression may have prohibited a wider use of the phonograph in the 1930s. A definite discouragement to its use was the Carnegie-funded Coleman report of 1929, which stated that the reading skill should be emphasized (Parker, 1961). Nevertheless, it should be noted that the decade saw much interest in the use of radio for foreign-language instruction. From October 1935 (volume 20) through December 1946 (volume 30), the *Modern Language Journal* had a radio "department."

It is not until 1908 that there is any evidence of a laboratory arrangement of phonographic equipment (Léon, 1962). By this is meant a dedicated facility for foreign-language study. This lab was at the University of Grenoble in France. An American, Frank C. Chalfant, who studied there in the summer of 1909, appears to have been the one who brought the idea back to this country. He installed a "phonetics laboratory" at Washington State College in Pullman during the 1911–1912 academic year. Pictures of this installation in use show students listening via networked earphones. This lab also had a phonograph recording machine so that students could compare their pronunciation with the native-speaker models.

Near the time that Chalfant established his phonetics laboratory, the U.S. Military and Naval Academy set aside rooms for listening to foreign-language records (Clarke, 1918). Another early facility was set up at the University of Utah in 1919 by Ralph Waltz (1930). He moved to Ohio

State and built another lab about which he published several articles (Waltz, 1930, 1931, 1932). Waltz is usually credited with coining the term *language laboratory* in 1930 (Hocking, 1964). In fact, Chalfant had used it as early as 1916 in the Washington State College yearbook, the *Chinook,* and probably in the regional foreign-language education circles of which he was a leader. In any event, the preferred term until after WWII was "phonetics lab." During the 1930s many institutions established labs (Gullette, 1932), but, as in the case of the phonograph, discussions of their use did not loom large in the methodological literature. For example, the *Modern Language Journal*'s annual annotated bibliography of monographs and articles only had four entries prior to 1945 besides the three articles by Waltz. The bibliography of the language laboratory for the years 1938–1958 compiled by Sanchez (1959) only added four items to the total for the prewar period. Eddy (1944) and Whitehouse (1945) describe the equipment and use of two labs at the end of the period under consideration.

28.8.1.2. History of the Use of Audio Resources in Foreign-Language Education: 1946 to 1958. The year 1946 is considered to mark the beginning of the modern language laboratory movement (Hocking, 1964; Koekkoek, 1959). The labs at Louisiana State University (Hocking, 1964) and the University of Laval in Quebec City, Canada (Kelly, 1969), were built that year. Whether these postwar labs owed anything to the previous phonetics labs is unclear, but probable. Claudel's (1968) use of "predecessor" (p. 221) expresses linkage. However, according to Koekkoek, "the beginning of the language laboratory movement was a new start, albeit with similar means and ends, rather than a direct expansion of the limited phonetics laboratory tradition" (1959, p. 4). Sanchez (1959) is ambiguous on the question. The earliest entry in his annotated bibliography of the "modern" language laboratory is a reference to a phonetics laboratory (Peebles, 1938), but he included the note "not related to the modern language lab, as such" (p. 231). The record at the universities of Iowa (Funke, 1949) and Tennessee (Stiefel, 1952) indicate continuity with phonetics labs. It thus appears that Koekkoek's statement must be tempered. Most institutions that built language labs after the war did so for the first time, whereas a few others updated their prewar phonetics labs.

A point of difference between phonetics labs and language labs were individual booths or carrels. Although the lab at Ohio State had long tables divided into "compartments" (Waltz, 1930, p. 28) by 18-inch-tall boards, these did not provide sufficient acoustic isolation (Schenk, 1930). Levin (1931) suggested that the facility he described would be improved by the installation of soundproof booths. These became standard equipment in the postwar labs. Labs of the period were principally audio installations, but movie, slide, and filmstrip projectors were sometimes present as well (Newmark, 1948; Hirsch, 1954).

Also at issue is the impulse for the modern lab movement. It is certain that the military's success in language training during the war caught the attention of the foreign-language teaching profession at large. In 1945, the *Modern Language Journal*'s annual bibliography began a separate category for the "Army" (Army Specialized Training Program, ASTP) method. It contained far more entries than any of the other 21 categories. Regarding labs specifically, Gaudin (1946) demonstrated enthusiasm for the Army experience in her design of a facility. Koekkoek maintained that "the language laboratory and its spread is a postwar development, fostered by a climate of experimentation which was stimulated by the Army language teaching program during the war" (1959, p. 4). Pictures of labs in the 1950s certainly have a military air to them. Students sitting in rows with eyes straight ahead suggests columns of soldiers at attention. The individual student in a booth wearing a headset evokes an image of a navigator or radar technician at his post on a ship or airplane.

Hocking (1964), however, adamantly denied that the ASTP method drove the establishing of labs. He is echoed by Barrutia:

> . . . we have Elton Hocking to thank for almost singlehandedly trying to keep the record straight about the fiction of the supposed extended use of recording equipment and aural-oral techniques in the A.S.T.P. . . . the Army Specialized Training Program did not, as is so widely believed, pioneer language laboratories . . . (1967, p. 890).

To what then did Hocking and Barrutia attribute the postwar interest in labs? They cite the availability of magnetic tape and tape-recording machines from 1946. Hitherto, labs were outfitted with phonographs or wire recorders. These had several problems: Their sound fidelity was low, they were fragile, and they were difficult to edit. Plastic disc player/recorders such as the SoundScriber (first advertised in the *Modern Language Journal* in October 1946) were an improvement, but according to Hocking "the superiority of the tape recorder-reproducer was immediately apparent" (1964, p. 18), and the verdict of history confirmed this first impression.

This major technological improvement does not fully account for the language laboratory movement. Roughly concurrent with the invention of magnetic tape was the development of the audiolingual method. It is here that the ASTP can be given some deserved credit. It stressed the listening and speaking skills more than reading and writing—the priorities of prewar methods. The Army method relied much on small-group practice to develop the learners' aural and oral abilities. Another important feature of the ASTP was the preponderate use of native-speaker instructors. Stack connects these developments in equipment and methodology:

> The language laboratory owes its existence to the recognition that the spoken form of language is central to effective communication, and that it should have as large a share in instruction as do written forms. In order to implement this new orientation of language teaching, the textbook (which is essentially graphic) was supplemented by sound recordings of native speakers. The coincidental advent of the tape recorder created a fortuitous juncture of technology and pedagogy (1971, p. 3).

By 1958, in the United States there were 64 labs in secondary schools and 240 in colleges and universities (Johnston & Seerley, 1960). The passage of the National Defense Education Act that year ushered in a new phase in language laboratory history.

28.8.1.3. History of the Use of Audio Resources in Foreign-Language Education: 1959 to Present. Responding to the challenge to Yankee know-how and American ingenuity by the Soviet's launching of Sputnik on October 4, 1957, Congress passed the National Defense Education Act, which President Eisenhower signed on September 2, 1958. The act sought to strengthen the teaching of mathematics, science, and foreign languages in America's schools. The intent of the foreign-language provisions of this important legislation has been described by Derthick (1959). The history of the language laboratory in the first years following the NDEA has been written by Parker (1961), Hocking (1964), and Diekhoff (1965). There was an explosion in the number of facilities, thanks to generous federal support: $76 million in matching funds by 1963 (Diekhoff, 1965). By 1962 there were approximately 5,000 installations in secondary schools (Hocking, 1964). Another 1,000 secondary schools had labs by 1964 (Diekhoff, 1965). This represents a thousandfold increase in the number of labs at the secondary level from 1958! Most of these were in medium-to-large school districts (Godfrey, 1967). Although colleges and universities were not eligible for equipment funds under the NDEA, they were caught up in the national enthusiasm for language study, and thus committed their own monies to labs. By 1962 there were 900 labs in higher education (Hocking, 1964). More postsecondary labs were built from 1965 when matching funds became available under Title VI-A of the Higher Education Act (Ek, 1974).

Unquestionably, the 1960s were the golden years of the language laboratory. Those involved in these facilities felt an urgent need to gather and compare experiences. William Riley Parker wrote this about the motivation for the first of the Indiana and Purdue universities–sponsored language laboratory conferences in 1960 (the others were in 1961, 1962, and 1965):

> . . . foreign language teachers feel themselves suddenly involved in a technological revolution, suddenly chin-deep in a tide of new demands upon their competencies, and they seek, some almost frantically, enlightenment and practical help (1960, p. v).

B. F. Skinner spoke at the first of these meetings (January 22, 1960) on the use of teaching machines for foreign-language instruction. One of the respondents to Skinner's paper was Robert Glaser. Neither of these men were foreign-language educators by training, but both were already well known in the educational technology community. Their presence at this conference is testimony both to the interest the larger educational community of the day had in foreign-language instruction and to the willingness of foreign-language professionals to accept insights from other disciplines, notably psychology. In addition to the Indiana conferences, there were many lab-related presentations at meetings of the various professional associations to which language educators belonged: the Modern Language Association (MLA), the American Association of Teachers of French (AATF), the American Association of Teachers of German (AATG), and the American Association of Teachers of Spanish and Portuguese (AATSP).

A spate of books, articles, and dissertations also accompanied the flow of money and the many installations. Most of the entries in Davison's (1973) 780-item bibliography of the language laboratory from 1950 through 1972 are from the 1960s, and thus post-NDEA. The first edition of Edward Stack's textbook, *The Language Laboratory and Modern Language Teaching,* appeared in 1960. It should be consulted by those interested in the literature of the period, because it explains the terminology of installations and operations current at the time. Also appearing in the early 1960s were Hutchinson's monograph concerning labs in high schools (1961), and the technical guide to facilities by Hayes (1963). Léon's book *Laboratoire des Langues et Correction Phonétique* (1962), although written in French and published in France, circulated widely in this country, as evidenced by the numerous citations of it. The Scherer and Wertheimer (1964) book-length report of an experiment involving language labs will be discussed in the section on research. As for articles, hundreds appeared in all ranges of periodicals from school district newsletters to long-established refereed journals such as *The Modern Language Journal, Language Learning, Hispania, The French Review,* and *The German Quarterly.* A publication that focused on language laboratories, *The Audio-Visual Language Journal,* was founded in Great Britain in 1962. Both *The International Review of Applied Linguistics* and *Foreign Language Annals* carried articles about the language laboratory from their inceptions in 1963 and 1967, respectively. The major research articles of the period will be noted in a later section.

During the first half of the decade, language laboratory professionals held caucuses at the conventions of the Modern Language Association and the Department of Audiovisual Instruction, but they soon felt the need for their own organization. The National Association of Language Laboratory Directors (NALLD) was founded in 1965. The NALLD began publishing a newsletter the following year. The inaugural issue reported that at the first NALLD meeting in Chicago in December 1965, there had been much discussion of the lab director's job description and the problem schools face in recruiting qualified applicants. Job openings were featured regularly from the start of this publication.

The major technical development of note during the decade was the audiocassette (Dodge, 1968). The advantages of this were the lower price and smaller, lighter machines that could play it. However, it did have the drawbacks of lower fidelity and greater difficulty of editing by cutting and splicing. The quality of sound was eventually ameliorated, and the editing problem was not sufficient to prevent the cassette from replacing reel tape in language labs in the 1970s. Another technical advance was the speech

compressor-expander (see 28.5). This device allowed a recording to be sped up (compressed) or slowed down (expanded). Articles on this technology were numerous in the general educational literature from the start of the decade. Paradoxically, it was not until 1978 that anything on it appeared in the NALLD Journal (Harvey, 1978). One would have expected a greater enthusiasm among the foreign-language community. The ability to slow down a tape would seem to be a boon to students struggling with a difficult passage. Moreover, variable-speed technology was not unknown in foreign-language teaching, for Hirsch (1954) had commended the use of the *sound stretcher* (p. 22) in the early 1950s.

Machines with a repeat or skip-back function came on the scene at this time as well. This feature permitted students to replay easily a tape segment, and thus was well suited to dictations and audiolingual listen-and-repeat drills. The cassette Canon Repeat-Corder L was first advertised in the *NALLD Journal* in the October 1970 issue. Aikens and Ross (1977) wrote an article in the same journal describing a reel-to-reel machine they fabricated. By the end of the decade, the major manufacturers, such as Sony and Tandberg, were producing machines with skip-back capability.

Language laboratories ended the 1960s on a mixed note. On the one hand, federal funding was diminished:

> . . . the amount of equipment funding in Title III-A of the National Defense Education Act (NDEA) and Title VI-A of the Higher Education Act (HEA), two large sources for equipment funds, dropped from an allotment in fiscal year 1968–69 of $91.24 million to nothing in fiscal year 1969–70. The portent of this budgetary reduction is not as black as it might seem: any program for which the federal government is still offering subsidy, e.g., bilingualism, poverty, etc., still has access to equipment funds, but the inflated years of the mid-sixties have come to a close (Dodge, 1968, p. 331).

On the other hand, Smith (1970) did not lament this decline in federal support as entirely negative, because he candidly acknowledged that "the recent years have seen much professional neglect and misuse of the language laboratory (p. 191). He sensed a positive development in the unanimous agreement that the laboratories should be used to "individualize instruction," in the university community and provide the corresponding "increase in expenditures for equipment and materials for tutorial and individualized instruction" (p. 192). Heinich (1968) also commented on the problems associated with labs and the insights that were gained by both language educators and instructional technologists:

> The *language laboratory* movement threw content and media specialists together in an intimate working relationship that produced very strange and startling experiences. For the first time, language teachers discovered that the mode and materials of instruction interact with instructional behavioral objectives and methods. Many language teachers did not understand that a language laboratory requires a different method of instruction: that print stimulus methods are not audio stimulus methods. On the other hand, the audiovisual specialist was shaken out of a comfortable bookkeeping-procurement function and introduced, often for the first time, to the rigors of developing curriculum materials to meet specific curricular objectives. The novelty of the roles played by both has caused so many difficulties that the language laboratory has not yet reached its potential value. One of the lessons learned by audiovisual directors in this encounter is the incredible quantity of materials required by technology when media are used for direct instruction. The classroom teacher, at the same time, was experiencing another instance of shared responsibility with media (pp. 50–51).

The 1970s and early 1980s were a period of malaise for the language laboratory. Coinciding with the drying up of funds was a sharp dropoff in the number of articles published. An index of this change can be seen in the ACTFL yearbooks. The first two volumes contained the articles by Dodge (1968) and Smith (1970), with 84 and 95 citations, respectively. The 1971 volume had one paragraph about labs and two references! From then on until 1983, many volumes contained no mention of labs, and those that did accorded a page at most. Labs had their vocal defenders to be sure (Jarlett, 1971), but frank avowals of their problems (Altamura, 1970; Racle, 1976) and their need for revitalization (Strei, 1977) were prominent.

A turnaround in the decline of the language lab could be seen from the early 1980s. A 3-day colloquium with the theme "A Renaissance for the Language Lab" was held at Concordia University in July of 1981 (Kenner, 1981). McCoy and Weible maintained that the recent "revival of interest in language laboratories" was "directly attributable to the 'domestication' of the tape recorder, made possible through the invention of the audiocassette" (1983, p. 110). What this indicates is that it took nearly 2 decades for the audiocassette to fully work its way into the instructional mores of teachers.

The lab of the 80s was not to be limited to audio technology. The year 1983 saw the founding of the Computer Assisted Learning and Instruction Consortium (CALICO), a group that was (and still is) dominated by language educators. It should not be thought that the invention of the personal computer in the late 70s was solely responsible for the interest in computer-assisted language instruction (known as CALL). Mainframes had already been much used for this purpose, most notably in the PLATO system at the University of Illinois. Computers were welcomed for their potential, but cautions were issued about the need to avoid the unrealistic expectations associated with early language labs and the need to learn other lessons from language lab history (Marty, 1981; McCoy & Weibel, 1983; Lemon, 1986; Pederson, 1987; Otto, 1989).

Ely's *Bring the Lab Back to Life* was published in 1984. In 1985 the president of the International Association of Learning Laboratories (IALL, the new name for the NALLD as of November 1982), Glyn Holmes, could affirm that the professional group was showing new signs of *vitality* (Holmes, 1985, p. 5). This rebirth was also indicated by volumes 18 and 19 of the ACTFL Foreign Language Edu-

cation Series, which were devoted entirely to technology (Smith, 1987, 1989). With new life came a new look. By 1989, Otto could write that "language laboratories have been redefined as multimedia learning centers that deliver computer and video services to faculty and students in addition to familiar audio resources" (1989, p. 38). Since 1988 the reinvigorated IALL has published several monographs dealing with learning-center design and pedagogical use (e.g., Stone, 1988) and has produced several "video tours" of facilities around the country.

A further sign of the broadening of focus of language laboratories is the greater attention placed on reading and writing. A prime example of the former is the popular series of software produced by the Transparent Language Company. The Systeme-D writing assistant program of Heinle & Heinle Publishers is an indication of the latter. Major research has accompanied its usage (Bland et al., 1990).

The use of video is firmly established in foreign-language teaching. A prominent instance is the innovative first- and second-year French course that appeared in 1987, French in Action. Interestingly, an early leader in the post-NDEA labs, Pierre Capretz, was the driving force behind it. French in Action and its subsequent Spanish equivalent, Destinos, received major funding from the Annenberg Foundation and are broadcast on many Public Broadcasting System stations.

28.8.1.4. Conclusion. Surely language laboratories represent the single largest investment and installment of audio resources in education. It is no accident that the foreign-language teaching community has been heavily involved in using audio. Audio has face validity in foreign-language instruction simply because much of language use is oral/aural. Granted, there has been concern that the reading and writing skills might be neglected in methodologies that make much use of recordings such as audiolingualism. Nevertheless, for foreign-language educators it has never been an issue of whether to use audio; it has been a question of how.

28.8.2 Research

The preceding historical account detailed the growth and extent of a particular application of audio technology. What has not yet been assessed is the effectiveness of this massive expenditure of effort and money. This is the task of research. This section will give the main currents of research for each period in the language laboratory's history. Details of each study will not be mentioned except insofar as they are crucial to interpreting the chief findings. The bibliography will permit the interested reader to locate and directly consult the reports cited for further information about the design and conditions of each study.

28.8.2.1. Research on the Use of Audio Resources in Foreign-Language Instruction: 1877 to 1946. There appears to have been very little attempt to provide an empirical justification for the use of the phonograph and phonetics laboratories before World War II. This is not entirely surprising, given that before the 1960s very few foreign-language scholars had training in quantitative experimental techniques: They were humanists schooled in literary and philological research methods. There are, however, accounts of problems with the use of phonographs and phonetics labs which can perhaps be classified as observational research. These observations will be noted, for they raise issues that were to be examined more rigorously later. Moreover, these records demonstrate that there was some notion of accountability among those who used early audio resources. That is, the phonograph and phonetics labs were not accepted and used uncritically.

Based on his *long experimentation,* C. C. Clarke (1918, p. 120) provided the first guidelines to appear in the scholarly literature on the proper use of the phonograph in teaching foreign languages. He granted that some teachers found the *mechanism* (p. 122) troublesome, time-consuming, and distracting. To this he countered that it afforded learners the opportunity to hear consistent native-speaker models that never suffered fatigue. He concluded that "the true success of the speech record is in teaching pronunciation and that nothing else should be expected of it" (p. 120). The emphasis on pronunciation training certainly became the hallmark of the phonetics laboratories. Waltz, the founder of the lab at Ohio State University, also cited the benefit of having tireless native-speaker models to imitate. By having the "constant control sounding in his ears" (p. 29), the student could exclude the imperfect approximations of his peers and gain confidence in his own speaking ability. However, a colleague of Waltz, Emma Schenk, complained that the earphones did not adequately keep out others' voices (1930). In addition, she complained of the poor audio quality and the lack of supervision in the Ohio State lab. She worried that students would "cultivate errors" (p. 30). She also noted much cheating on time slips and many students who were not on task while in the lab. Levin (1931) was sympathetic to labs and sought to offer constructive criticism of their use. He stressed the need for immediate feedback so as to avoid the problem Schenk had feared, namely, the development of bad speech habits. Gullette (1932) showed that this fear was justified. He noted with consternation that many students working alone in the lab reverted back to the poor pronunciation practices that earlier had been eradicated in class drill sessions. He stressed that imitation was not sufficient; what was needed was ear training such as was done in music classes. This would allow for self-diagnosis and correction.

Waltz's report (1932) of two studies he consulted on, but did not conduct himself, is the first record of an attempt to establish empirically the phonetic/language laboratory's effectiveness. It is ironic, in view of the identification of the language laboratory with foreign languages, that neither investigation involved their teaching! The first experiment had to do with the teaching of the Irish accent; the second was concerned with correct English diction. Both studies can be faulted for the low number of subjects (20 and 24),

the apparent nonrandom assignment of subjects to treatments, and the lack of statistical analysis beyond a comparison of group means. Nevertheless, Waltz did note that the groups were equivalent by using scores on standardized tests of intelligence, hearing, and pitch discrimination. In the first study, the lab group's mean was 10.1 (out of a possible 20 points). The control group's mean was 8.04. In the second study, both the lab and nonlab groups showed similar gains. Waltz argued that the comparable improvement was actually evidence in favor of the efficiency of the lab: Class and instructor time was saved by having students work independently in the lab.

For the sake of comprehensiveness, Peebles' master's thesis (1938) must be mentioned. It was included in the annotated bibliography compiled by Sanchez (1959). Students who volunteered to use the Phonetics Laboratory at the University of Colorado and who received one or two French pronunciation tutorial sessions were compared with students who did not avail themselves of these opportunities. Amazingly, she did not specify how much the volunteers used the lab. Neither were the total number of subjects, nor the number of subjects per group, specified. These omissions bespeak a blatant lack of control that invalidates any conclusions that might be drawn from her data, which in fact consisted only of mean numbers of pronunciation mistakes on a posttest.

28.8.2.1.1. Summary. Obviously, no firm conclusions can be drawn about the effectiveness of the phonograph and the prewar phonetics laboratory from these few observations and two cursory investigations. There appears to have been a consensus among practitioners that the best use of this equipment was for pronunciation training. All saw a potential benefit in untiring, consistent, native speaker models for students to imitate. However, complaints were raised about the sound quality of recordings, and it was observed that many learners lacked the self-monitoring ability to profit fully from them. Just as the next period of language laboratory history saw an increase in the number and sophistication of facilities, so there was similar growth in the inquiries concerning their value.

28.8.2.2. Research on the Use of Audio Resources in Foreign-Language Instruction: 1946 to 1958. Language laboratory research of the postwar and pre-NDEA period may be described as nascent. Certain features of empirical research are seen; some are only partially present, and others are completely absent. For example, one sees the first use of standardized tests as criterion measures, and this use is universal. On the other hand, only one study (Allen, 1960) randomly assigned subjects to treatments; intact classes were used otherwise. Only two-group designs and *t* tests were used. The number of subjects, when reported, was uniformly low. There certainly was not an agreed-upon research agenda. In fact, researchers of the day were either unaware of what their peers were doing (there is little citation of others' work) or they simply ignored it. With these limitations in view, the following discussion will list five studies of the period in chronological order and present their conclusions. According to Kelly (1969), more experiments were conducted than this number would suggest, "but we only know of those whose authors had the time and energy to write articles about them" (p. 245). This is corroborated by Johnson and Seerley (1960), who refer to studies done at a high school and two universities (all unnamed) and of research that was planned at the University of Massachusetts.

Stiefel's description (1952) of the language laboratory at the University of Tennessee and its usage is barely beyond the anecdotal level. Yet its mention of the University of Chicago language investigation tests and the cooperative tests (created by the forerunner of Educational Testing Services) does represent the first, inchoate desire of those involved in language labs to have an objective benchmark with which to compare groups of learners who used the lab with those who did not. In this case, Stiefel compared the scores of lab classes on these measures and on an in-house test with classes from previous years. Thus, this is an ex post facto study. He noted higher scores for lab groups on the in-house tests, but he was hesitant to draw any strong conclusions from these. He found that both groups were comparable on the standardized tests. This he took as heartening evidence that the reading ability (as measured by the cooperative test) of the lab groups did not suffer because of their emphasis on the listening and speaking skills. This last point was of great concern to the scholarly community of the day, as further evidenced by the following study.

Supported by a grant from the Carnegie Foundation for the Advancement of Teaching, Brushwood and Polmantier (1953) at the University of Missouri sought to determine whether dialogue repetition and memorization in the language lab increased learners' aural skills. Although for administrative reasons they were unable to randomly assign subjects to treatments, these researchers did take the trouble to administer the Iowa Foreign Language Aptitude Test to the intact classes that constituted the treatment groups. Moreover, the researchers obtained access to the scores on two English proficiency tests that all the subjects had taken previously. All these tests revealed that the control and experimental groups were matched on these measures, as they were in age.

Four groups were formed: two groups of 19 subjects each who were enrolled in elementary Spanish, and two groups of 23 who were enrolled in elementary French. The control groups simply attended the standard 5-hour per week (1 hour daily) course as taught at the University of Missouri. The experimental groups covered the same material (grammar, reading, and composition) as the control groups, but did so in 4 hours instead of 5. The experimental groups also attended two 1-hour laboratory sessions during the first 4 days of the week. In these sessions, they worked with a dialogue written for the experiment that incorporated the grammar and vocabulary that was studied that week. The work consisted of listening to the dialogue via earphones and chorally repeating it until it was memorized. A graduate student or upperclassman lab attendant controlled the tape

player and thus directed the sessions. His or her only other task was to correct gross pronunciation errors. The experimental group then had a fifth class session in which the regular instructor had the students review and act out the dialogues. The dialogue was then manipulated by changing number, person, tense, object, etc., as a transition to free conversation. This fifth hour was deemed "the crucial point in the achievement of the oral-aural objective" (p. 8).

At the end of the semester the groups were given the cooperative tests on reading, vocabulary, and grammar, and an aural comprehension test created for the experiment. For whatever reason, both *t* tests and *F* tests were calculated for the two Spanish and two French groups, but no tests were run on a combination of control and experimental groups across languages. The results showed that there were no significant differences on the cooperative measures. There were significant *t*s, but not *F*s, in favor of the experimental groups on the aural comprehension test.

This study can be faulted on several grounds, but perhaps the most serious flaw may be the lack of control for amount of instruction. Although the authors claimed that the 2 hours of lab practice for the experimental groups were in lieu of homework required of the students in the control groups, it must be noted that the lab sessions were scheduled and monitored. Whether students in the control sections did their work or not is unknown. Moreover, the significant difference between the groups on aural comprehension was measured by a nonstandardized test, the validity and reliability of which is open to question. All of these criticisms aside, Brushwood and Polmentier's study was certainly more rigorous than previous investigations of the use of audio resources in foreign-language teaching.

Next in chronological order are two ex post facto studies that are included here for the sake of completeness. The first is the description by Fotos (1955) of the use of the language laboratory at Purdue University. In direct opposition to the Brushwood and Polmentier study, the lab at Purdue was used for "*predrilling* [emphasis added] the student on the French text of the basic grammar or reading lesson" (p. 142). Fotos reported that students in first-year French scored 60.1 on the cooperative tests; second-year students scored 71.3. The national averages were 56.7 and 68.8, respectively. Whether this was a significant difference cannot be ascertained.

Mueller and Borglum (1956) looked at correlations between lab attendance and course grade, final exam score, and cooperative test score at Wayne University. They noted that students who voluntarily attended the lab more than the minimum requirement of 30 minutes per week generally did better on these measures. They drew special attention to the heavy lab users' 10% increase on the cooperative reading test: "an unprecedented jump in 8 years of recorded scores" (p. 325). Moreover, they observed that even students who only attended the lab 30 minutes per week scored better than students from previous years who had no lab experience. They also noted a lower drop rate for heavy lab users. One can surmise that greater time-on-task naturally produced greater learning. In their discussion, Mueller and Borglum also acknowledged a significant teacher effect: The lab's director "succeeded in getting the students of his sections to attend the laboratory 2 or 3 times more frequently than other instructors" (p. 322).

Allen (1960) conducted a study during the 1957–58 academic year which represents the last investigation of language laboratories in the 1946–58 period. The 54 subjects were 15- and 16-year-old students in a high school operated by Ohio State University. Allen created eight groups based on level (elementary or intermediate), language (French or Spanish), and use of the lab (55 minutes per week or none). These divisions made for groups as small as five. He administered three standardized tests in order to have a basis for pairing subjects. Once the pairs were established, he used a random-choice technique to assign students to the lab or nonlab treatments.

The lab groups spent one classroom hour listening to instructor-made tapes of "humorous or suspenseful tales" (p. 355) and answering questions about them in the target language. They recorded their answers and then spent the rest of the period listening to commercially prepared recordings. There was absolutely no written material presented during the lab hour. The nonlab group read the same stories and answered the questions in writing. If any time remained, they did free reading from a collection of books at their level.

At the end of the school year, all groups were given three standardized tests (including the cooperative) that measured reading, vocabulary, grammar, speaking, and listening. Allen only reports means and standard deviations. In all cases except one, the laboratory groups scored identical to or higher than the nonlab groups. The exception was the Intermediate Spanish lab group ($n = 5$), which scored lower on the speaking test. In several cases, the differences between the means were large, but Allen did not compute any test of significance. In his brief conclusion, however, he claimed that the laboratory groups "achieved significantly higher scores in reading, vocabulary, and grammar" (p. 357), but that there were no differences in speaking or listening. This author calculated a *t* test on the cooperative French test means for the largest groups, those in Elementary French ($n = 10$ each). The lab group had a mean of 57 (s.d. = 23); the nonlab group mean was 39.4 (s.d. = 20). This turned out to be significant at the 0.001 level.

It is fitting that the last of the studies of the 1946–58 period should be the one with the highest methodological standards. Yet the number of subjects was quite low for the design chosen, and it is baffling that Allen claimed to have found a significant difference in favor of the lab groups, but did not bother to report any data beyond means and standard deviations. Moreover, it is ironic that reading, grammar, and vocabulary scores were enhanced by listening in the language laboratory, whereas listening scores proper did not reveal any difference between the lab and nonlab groups. Thus, Allen's study gives weak but curious evidence of the language laboratory's contribution to foreign-language learning.

28.8.2.3. Summary. Writing in the early 1960s, Carroll (1963) stated that virtually all previous foreign-language research "has only rarely been adequate with respect to research methodology" (p. 1094). For him, language laboratory research was not an exception to this rule. He briefly reviewed three studies concerning labs; these were not included in this section because they did not contain important results, were not widely circulated at the time (two were institutional reports), and were not cited by subsequent researchers. Therefore, what one can conclude from Carroll's review and this summary is that while the research during the 1946–1958 period did not firmly establish the positive value of language laboratories, it did provide circumstantial, and in one case (Allen, 1960) empirical, evidence in favor of this conclusion.

Writing at the close of the period under consideration, Koekkoek (1959) stated that labs were so "firmly established" in language teaching that "no teacher can remain today unaffected and disengaged" (p. 5). He went on to describe the ambivalence about them within the profession and closed his article with the hope that subsequent experience would resolve "basic questions to be expected from the use of laboratory machines and the best methods of obtaining the results" (p. 5). If the nascent body of research could only offer a cautious "thumbs-up" assessment, it also showed that those promoting labs were willing to be held responsible for their use. This was fortunate, for during the next phase of the lab's existence, a period of great growth because of major expenditures, the public would eventually demand an accounting.

28.8.2.4. Research on the Use of Audio Resources in Foreign-Language Instruction: 1959 to the Present. The massive increase in the number of language laboratories, thanks to the NDEA, prompted a comparable increase in the amount of research concerning their effectiveness. In fact, some of the studies were funded by the NDEA under its Title VI provisions. The extent of this research is such that this section cannot detail every investigation that was undertaken. The several dissertations listed by Davison (1973) will not be treated. This discussion will focus on four large-scale studies of labs: three in high schools and one in a university. These all received much attention at the time. Moreover, those studies that have been thoroughly reviewed elsewhere will be only briefly described.

During the 1961–62 school year, Keating (1963) conducted a study of the use of the language laboratory in French classes in New York City high schools. He cited Allen's study (1960) as the "only exception" (he was evidently unaware of the Brushwood & Polmantier study) to the rule that "the literature abounds with articles that describe the benefits of using language laboratories" but "contains virtually no reports upon the empirical validation" (p. 8) of them. He called Allen's results "quite interesting" but noted a possible Hawthorne effect, which he felt "severely compromised" (p. 8) them. Keating knew of the research being simultaneously conducted in New York City by Lorge (to be described later).

Keating's was a large-scale study involving approximately 5,000 subjects in 21 school districts. Schools were divided between laboratory and nonlaboratory users based on a questionnaire filled out by each district's foreign-language coordinator. Besides this factor, groups were formed according to year of study (first through fourth years) and IQ scores (five levels). The dependent measures were reading comprehension, listening comprehension, and speech production. The cooperative test was used to test the first two skills; however, first-year students were not given the listening portion because it was designed for intermediate and advanced students. The French speech production test was used to evaluate speaking. This instrument was constructed specifically for the study. Of note is that it was not administered to all subjects: only 519 students from 12 of the participating school districts were given it. The results showed a sole significant finding in favor of the lab groups, on speaking among first-year students. Otherwise, there were several cases of the nonlab groups scoring significantly higher.

Keating's findings were promptly and vehemently disputed. The April 1964 issue of the *Modern Language Journal* included four rebuttals (Anderson, Grittner, Porter & Porter, Stack). The criticisms showed much overlap. Keating was taken to task for numerous methodological flaws: failure to define what was meant by language laboratory and the activities that went on there, failure to control for amount of time spent in the lab, failure to control for the socioeconomic level of the schools and the quality of their lab installations, use of *t* tests when ANOVAs were called for, and sloppy reporting of results (the number of subjects per group was not consistent). Keating was also criticized for using several different IQ tests, rather than one, to group subjects. The validity of his speaking test was challenged for being in fact only a pronunciation measure. Keating was shown no mercy: Despite the disclaimers he gave about the generalizibility of his results, he was accused of spreading anti-lab propaganda by Grittner.

Because the literature of the period contains no defense of Keating's study, it can be concluded that it was dismissed by the scholarly community of the day. Unfortunately, the public was of another mind. It seized on the notion that if language laboratories are not useful, then the massive investment of tax dollars in facilities was a waste. An example of this attitude was a newspaper editorial about the Keating study entitled "Backwards Via 'Aid'" that was reprinted in the *Modern Language Journal* issue containing the four rebuttals. Such a response gives credence to the propaganda charge made by Grittner. He and Stack and Anderson pointed out, with great dismay, that the Institute of Administrative Research of Columbia Teacher's College, which had sponsored Keating's study, mailed out a five-page preliminary report to school administrators across the country. They viewed such an action as unprofessional; it was clearly inflammatory in its impact.

Lorge (1964) conducted two experiments in New York City high schools. The first took place during the 1961–62

school year, and the second was done the following year. Thus, the first study coincided with Keating's investigation. Whether there was any overlap of subjects between the two studies is unknown, but could hardly be problematic given that only two schools were involved in Lorge's first inquiry; Keating's entailed 21 districts. Lorge described the purpose of her study thus:

> The object of the study was not to compare what a student learns from a teacher alone as opposed to what he learns from laboratory work alone. The question was whether the teacher improves the teaching-learning situation by using the laboratory as a teaching aid. The research was intended not to give the laboratory a passing or failing mark—if it passes, use it; if it fails, rip it out—but rather to determine in which areas it had proved to be successful, and how its use could be made more effective (p. 409).

The first study compared first-, second-, and third-year French classes. Unfortunately, the number of classes and subjects is not specified in the article, and the full report of the study is not available for consultation; by 1965 it was already out of print (Lorge, 1965). All that is known is that the classes were determined to be comparable based on the Stanford reading test and the Gallup-Thorndike vocabulary test. Half of the classes had 60 minutes a week of supervised lab practice in lieu of a fifth class period. The other half had five class meetings. The course content was the same for both groups. At the end of the school year, all classes were given the cooperative French test to gauge reading, vocabulary, and grammar skills. A speaking test and a listening test, both written by the experimenters, were also administered. All the tests contained subtests for which separate statistics were calculated. There were no differences between the groups on the cooperative test. The first- and second-year laboratory groups tested significantly higher than the control groups on the fluency component of the speaking test. The second-year laboratory group also scored significantly higher on the intonation component. The third-year laboratory group was significantly superior in listening.

The second experiment compared two types of laboratory equipment: audio-active and recording-playback. The first was a headset with earphones and a microphone; the second was an identical headset plus a tape recorder for each student. The other factor was time. Daily usage of 20 minutes was compared to a once-a-week 60-minute session. Five groups of second-year French students were formed. It should be stressed that none of the subjects had previous laboratory experience. Moreover, during the study, the control group did not use any equipment. The other four groups were formed by crossing equipment type and usage time. The dependent measures were the same as in the first study, with the addition of a mimicry test.

The *t* test results from the 14 components are difficult to interpret. Some differences are reported at a .01 level of significance, others at a .05 level, but it is impossible to determine whether one group was significantly higher than all the other groups or only some of them. The rankings that were also reported are more helpful, for they allow trends to be detected. On measures of enunciation, the order was thus: (1) daily record-playback, (2) daily audio-active, (3) weekly record-playback, (4) weekly audio-active, and (5) control. Thus greater time, frequency, and more elaborate equipment favor one aspect of the speaking skill. However, as regards lexical and syntactic features of speech, the control group was ranked first, with the daily record-playback group coming in second. This finding should be considered along with the result from the composite score on the cooperative test. Here, the daily record-playback group ranked first and the control group was second. The difference between the two groups was not significant, but both groups were significantly higher than the other three groups. What emerges is this: The daily record-playback group and the control group scored similarly, and significantly better than the other groups, on both oral and written measures of vocabulary and grammar.

From the above findings, one is tempted to draw an "all or nothing" conclusion: Either use a fully equipped lab daily or dispense with it altogether. It seems that certain outcomes will be the same in either case. The corollary is that infrequent usage of a modest lab actually appears to be detrimental to the lexical and syntactic aspects of language learning! However, Lorge does not make such a counter-intuitive deduction. She noted that in the first study, there were no differences between the lab and nonlab groups on the vocabulary and grammar tests. In the second study, she maintained that any measure showing statistically significant differences showed at least one laboratory group that equalled or exceeded the gains made by the control group. This appears to indicate that time spent in the laboratory contributes to conventional learnings as well as to listening and speaking skills (p. 419).

The last sentence is crucial. Taken together, these studies indicated an overall advantage for the language lab. Lorge also noted that a higher percentage of students in lab sections continued studying French beyond the 3 years required for high school graduation and college admission.

Lorge's study appears to have been well received by the scholarly community. Stack (1964) praised Lorge's work in his critique of the Keating study. Only Green (1965) ventured criticisms. Some of his complaints had to do with the manner in which the results were reported. He was more concerned with the apparent addition of another group after the study was underway. Lorge (1965) answered these objections easily in her rebuttal, which was included in the same issue of the *Modern Language Journal* as Green's piece.

In 1966, Philip D. Smith began an investigation of beginning high school French and German teaching and learning, which lasted through 1969. It was sponsored by the Federal Office of Education under Titles VI and VII of the NDEA and is commonly referred to in the literature as the Pennsylvania project because all the participating schools were in that state. Smith summarized his findings in 1969 articles in *Foreign Language Annals* (Smith, 1969a) and the *French Review* (Smith, 1969b), which are

more accessible than the technical reports he submitted as part of the grant's requirements. The October 1969 issue (volume 53, number 6) of the *Modern Language Journal* contained six articles critiquing the Pennsylvania studies. The December 1969 issue (volume 3, number 2) of *Foreign Language Annals* contained the summary article by Smith and two review articles. Contemporary synopses of the project and its reviews by D. L. Lange (1968) and W. F. Smith (1970) will be relied on for this discussion.

In the first year of the study, 2,171 students participated. Three teaching strategies and three language laboratory systems were compared. The strategies were: traditional, functional skills, and functional skills with grammar. By *traditional* was meant that an emphasis was placed on vocabulary acquisition, reading and writing skills, translation, and grammatical analysis. *Functional skills* was a synonym for the audiolingual method; the command of a core vocabulary and key syntactic patterns was emphasized, as were the speaking and listening skills. *Functional skills with grammar* was, as the name indicates, the addition of grammatical explanations to the audiolingual method. The three language laboratory systems were: audio-active, audio-active record, and tape recorder in the classroom. The first consisted of two, 25-minute practice sessions each week in which a 10-minute drill tape was played twice. The second arrangement differed from the first in that the students recorded their first practice with the tape and then listened to their own responses. Both of the audio-active groups also practiced in the classroom with a tape recorder each day under the supervision of the instructor for one-fifth of the period. The tape recorder in the classroom group did no lab practice. What they did was at least 10 minutes of guided practice with the tape each day in class.

The results from the first year indicated no significant differences between the teaching strategies, except for reading, where the traditional group outperformed the two audio-active groups. There were no significant differences detected between laboratory systems. During the second year of the project, 639 first-year students participated in a replication study, and 1,090 of the original 2,171 subjects were observed in their second year of language study. The results from this second year of the investigation were in line with those of the first. In the third, year the number of subjects (third-year students) dropped to 277, and by the fourth year it was down to 144 fourth-year students. The findings from these last 2 years showed the traditional students faring significantly better than the audio-active students in both reading and listening. In none of the 4 years of the study was a significant difference in outcomes found according to the laboratory system.

Although the Pennsylvania project generally received higher marks for its methodology than did the Keating report with which it was often compared, there were nevertheless several critiques leveled and questions raised. Some of these involved control issues, such as the degree of teacher adherence to experimental guidelines, the consistency of laboratory installations and maintenance between schools, and the lack of data as to the amount of time the labs where actually used. Carroll (1969b) detected stowaway variables and practice effects. Perhaps the most serious criticism was the claim (Valette, 1969) that the cooperative test was an inappropriate measure of listening achievement. It was maintained that the vocabulary in this test was closer to what was in the textbook used by traditional groups than the one used by the lab groups. Moreover, evidence from other sources was cited which indicated that the cooperative test was simply too difficult for students in their first 3 years of foreign-language study. This second criticism had broad implications: It cast doubt on the instrument that had been used in all previous language laboratory studies and in many other studies of foreign-language teaching.

Carroll (1969b) and Smith (1970) assessed the implications of the Pennsylvania project. For them, the supposed findings in favor of the traditional groups did not warrant a return to former means of teaching. Rather, they viewed the report, despite its faults, as a credible demonstration that the enthusiastic adoption of new approaches and accompanying materiel does not guarantee success. "The Pennsylvania studies have removed us from our tower of false security" (Smith, 1970, p. 208). For Carroll, the specific lessons to be learned were that audiolingual textbooks needed more linguistic content and that less emphasis should be placed on drills and other "habit formation" activities (1969; p. 235). Smith ended his review on an upbeat note: "It is time to meet the challenge of a new decade" (1970, p. 208). But such a positive attitude did not prevail. As was noted in the historical section above, language laboratories were in the doldrums in the 1970s and early 1980s. Davies (1982) singled out the Pennsylvania project for making complete the growing disillusionment of the period with labs. Moreover, it appears that the study discouraged other research, for it was the last of the large-scale inquiries into the language laboratory's effectiveness.

The only major inquiry of the language laboratory involving postsecondary students will now be discussed. Scherer and West (1964) described in a 246-page book the 2-year NDEA-sponsored investigation they conducted from September 1960. Their goal was to compare the audiolingual approach to the traditional grammar-reading method. Thus, this was not an examination of the language laboratory per se; rather, it was an inquiry similar to the Pennsylvania project (not yet conducted), which was interested in the language lab because of its intimate connection to the audiolingual method. The subjects were beginning German students at the University of Colorado. Intact classes were used, and these were determined to be similar on measures of general academic ability, language learning aptitude, and motivation, as well as sex, age, and year in school.

All of the teaching staff received a week of training in the respective methods prior to the start of the experiment. In addition, there were weekly meetings and frequent observations by the principal investigators and outside consultants to ensure that the instructors adhered to the experiment's guidelines. The traditional approach is only scantily

described, but the audiolingual procedures are elaborately detailed in Scherer and Wertheimer's book. The essence of the latter was dialogue memorization and related drill and practice in class. The frequency and duration of the lab sessions were unfortunately not specified; they were for "overlearning" (p. 83) the material presented in class. It is stated that the lab sessions were unmonitored and were of the "library-type" (p. 83), which presumably means the students attended at their convenience. Of note is the postponement of reading for the audiolingual group until the 12th week of the semester. To be specific, the audiolingual group saw absolutely no written German until that point. When reading began, it consisted of the dialogues that had been previously memorized and recombinations of the vocabulary contained in them.

The investigators claimed that they conducted a "persistent and continuous search" (p. 108) for standardized tests to use to measure the outcomes of the two teaching approaches. They were not satisfied with what they found, because "nothing that the major test distributors had to offer seemed to meet the requirements of our situation" (p. 108). They therefore constructed tests of the four language skills and two for translation: German-to-English and vice versa. The *t* test statistic was used for comparisons. At the end of the first year, the audiolingual students were significantly superior to the traditional students in speaking and listening. The superiority in speaking was maintained in the second year, but the advantage for listening was not. On the other hand, the traditional students significantly outperformed the audiolingual students on reading and writing during the first year, and maintained their edge on the latter skill during the second year. The traditional students also were higher in German-to-English translation during both years, and better in English-to-German translation in the first year.

In addition to these measures of linguistic proficiency, Scherer and Wertheimer also used standardized scales and questionnaires they constructed to evaluate the subjects' motivation to study German and their attitude to it and its speakers. They were also concerned with "habituated direct association." By this was meant the ability of the students to think in German, their inclination to translate or not, and their sensitivity to semantic nuances between the two languages. Numerous intercorrelations between these and measures of affective constructs such as anomie, social inhibition, and desire for further German study were calculated. The researchers summarized their work thus:

> The experiment has demonstrated that the two methods, while yielding occasionally strong and persisting differences in various aspects of proficiency in German, result in comparable overall proficiency. But the audiolingual method, whether its results are measured objectively or estimated by the students themselves, appears to produce more desirable attitudes and better habituated direct association (p. 245).

John B. Carroll (1969a) characterized the Scherer and Wertheimer study as "ambitious" (p. 869) and more rigorously designed than any previous examination of the audiolingual approach. He accepted the investigators' conclusions as valid, but offered the following:

> The conclusion that emerges from this experiment is that the differences between the audiolingual and traditional methods are primarily differences of objectives; not surprisingly, students learn whatever skills are emphasized in the instruction (pp. 869–870).

Besides the large-scale and well-publicized studies of Keating, Lorge, Smith, and Scherer and Wertheimer, there have been many smaller investigations since 1959. Eight studies that appeared in major journals have been selected for inclusion here according to chronological order. Only their main findings will be given, since these studies in general did not generate the interest of the larger studies that were described above.

Bauer (1964) found that university students who used the language laboratory in a supervised group-practice condition performed significantly better on oral and dictation measures, but not on a writing measure, than students who studied individually and were not supervised. Two drawbacks to the study were the low number of subjects ($N = 24$) and the use of nonstandardized tests. Moreover, a close examination of the data reveals that the supervised subjects as a group used the lab 125 minutes more over a 3.5-week period than the unsupervised subjects, so the observed differences could possibly be attributed to greater time-on-task.

Young and Choquette's NDEA-sponsored study (1965) was a series of seven experiments that sought to determine whether any of four language laboratory equipment configurations made a difference in the subjects' abilities to self-monitor their pronunciation. The systems were characterized by the feedback options they presented: (1) passive, (2) active, (3) long-delayed comparison, and (4) short-delayed comparison. The first three systems were standard options for language laboratory installations at the time. An apparatus for the fourth condition was specially fashioned for the study by the investigators. In the passive arrangement, the subjects repeated after taped prompts, but they could not clearly hear their responses because the headsets muffled their voices. In the active arrangement, subjects could hear their responses amplified through their headsets as they spoke. In third option, subjects could record their answers for later comparison. In the fourth setup, the students could hear their recorded response within 1½ seconds of making them. Subjects in the active feedback configuration were found to have slightly superior pronunciation than subjects in the other arrangements. However, the authors qualified this finding on several grounds. Of note was the lower sound quality of the fabricated equipment used in the short-delay condition. The authors admitted that this hampered a true comparison with the other three conditions.

Buka, Freeman, and Locke (1962) and Freeman and Buka (1965) conducted experiments that sought to establish psychoacoustic parameters for language laboratory equipment.

The first study determined that a high-frequency cutoff of less than 7,300 cps hindered subjects (high school students) from perceiving certain phonemic contrasts in German and French. The second study found that a low-frequency cutoff of 500 cps caused subjects (again high school students) to make significantly more errors in German phoneme discrimination than a 50-cps cutoff. However, no significant differences were found between these two levels for French phoneme discrimination. It was also found that consonant distinctions were more affected than vowel distinctions by the degradation of sound quality brought on by filtering.

Benathy and Jordan (1969) reported on a post hoc comparison of achievement scores in Bulgarian courses at the Defense Language Institute. The scores of 13 classes (87 students) that completed the course between August 1959 and September 1963 were compared to the scores of 15 classes (103 students) that finished between November 1963 and July 1967. The difference between these classes was the introduction in the fall of 1963 of the Classroom Laboratory Instructional System (CLIS): CLIS is a designed interaction of live instruction and a set of different kinds of learning experiences that make use of preprepared and recorded instructional materials, delivered through the electronic media (p. 473).

The authors stressed that the CLIS system kept the learners on task much more than in a typical classroom. This was because the earphones both isolated each learner from the erroneous responses and pronunciations of others and provided quality native-speaker models. Moreover, the learner did not wait to be called on as in a regular class; it was always his or her "turn." The equipment used appeared to be that of a typical audio-active language laboratory, although the authors do not use the term in their article. Curiously, they do not cite any language laboratory literature in their discussion, yet their description and justification for CLIS are identical to those commonly found in language laboratory writings.

The two groups were found to be very similar in ages and scores of the Army Language Aptitude Test. Class sizes were nearly identical, and the same textbooks and proficiency test were used throughout the 8-year period. It was found that the CLIS classes scored significantly higher than the pre-CLIS classes on the two skills measured by the test, namely, reading and listening. The differences were especially pronounced in the case of the latter skill.

Despite the many experimental controls and the marked differences between the groups, there are three questions that may be raised about this study. First of all, as no mention of instructors is made, one wonders whether teacher effects were held constant. Secondly, the generalizability of the results to high school and university students is doubtful, given that the subjects were all adults studying for specific career purposes at the Defense Language Institute. A third consideration is a question: Why did Benathy and Jordan not more fully report on the synchronous study that preceded the longitudinal one? They claimed similar significant results from it in favor of the CLIS. More information (i.e., number of subjects, a figure showing *t* values) about it would give greater credibility to their overall conclusion.

The Chomei and Houlihan (1970) study compared three language laboratory systems: instant playback, long-delay playback, and audio-active. The instant playback option allowed the subjects to have their recorded response to the program stimulus echoed back within half a second. The long-delay group had to rewind the tape to hear their recordings. The audio-active group did not record their responses. It can thus be seen that this study closely resembled what had been done by Young and Choquette (1965), but, surprisingly, this earlier work was not cited. The subjects in the Chomei and Houlihan investigation were 140 Japanese 10th-graders, who were all taught by the same instructor. It was found that the instant-playback group performed significantly better than the other groups on one out of five translation tests and on four out of five speaking tests that had been specially created for the experiment.

Sisson (1970) did a study that was sponsored by the U.S. Office of Education. Its aim was to settle the controversy among language educators as to the benefit (or lack thereof) of delayed comparison on students' ability to perceive and produce the phonemes of another language. Thus, this study shared the same goal as the work of Young and Choquette (1965) and Chomei and Houlihan (1970). That Sisson did not cite the latter is understandable, since it was contemporary to his own. What is surprising is that he ignored the former, yet did cite 39 other articles. In this oversight he followed Chomei and Houlihan, as pointed out before. Why a major study published in a leading journal was so ignored is an unanswered question in the record.

Sisson claimed that "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" (p. 82). The special equipment used in the study, the Plurilingua language laboratory, was thoroughly described. The subjects were 24 students of English as a second language at the University of Michigan. They were in three intact classes of eight students each. The classes were matched on the basis of a modified version of the test of Aural Perception for Latin American Students. This instrument had a phoneme discrimination section and two phoneme production portions.

Two conditions were compared. Half of the students (four from each of the three classes) listened to a taped stimulus and recorded their answer. On completion of an exercise, these subjects rewound the tape and repeated the exercise in the same manner. These subjects formed the "active group." The other group of subjects recorded their responses, as did the active group. However, at the completion of the exercise, these subjects rewound their tape and listened to their first responses rather than record them a second time. This was the "delayed-comparison group." Both groups spent 1 hour per week in the language laboratory during the 8-week term. The modified version of the test of Aural Perception for Latin American Students, which had been used as the pretest was also used as the

posttest. Sisson found no significant difference between the two groups on either discrimination or production.

Smith (1980) conducted a study to determine whether the slowing down of recorded material had a beneficial effect on listening comprehension. The reader will recall from the section on history that, during the 1960s, equipment became available which was capable of slowing down (expanding) or speeding up (compressing) recordings without distortion. Smith claimed that his search of the literature turned up no reference to studies addressing the specific application of this technology to foreign-language instruction. This claim was proved incorrect in Driscoll's article (1981), which listed several such studies. However, it should be pointed out that Driscoll was also guilty of oversight; he omitted Smith's study even though it was in the same outlet, the *NALLD Journal,* as his article.

Smith's subjects were second-semester students of French at West Chester State College in Pennsylvania. The control group had 11 members, and the experimental, 12. The cooperative test was administered as a pretest, and the control group was found to be significantly better in reading ability than the experimental group, but both groups were equal in listening comprehension, the skill at issue in the investigation. The study stretched over the fall 1978 semester. The control group covered 12 audio lessons that were recorded at normal speed. The experimental group listened to four lessons that were slowed by 20%, four that were slowed by 10%, and four that were at normal speed. At the end of semester, the students were again given the cooperative tests. Contrary to expectations, the ANCOVA and Finney *t* test procedures showed that the control group scored significantly higher on listening comprehension than the experimental group who listened to expanded material.

Despite such a clearcut albeit counterintuitive finding, Smith cautioned that the study needed to be replicated with a larger number of subjects and for other languages before it could be reasonably concluded that expanded speech was not beneficial, or perhaps even harmful, for the acquiring of listening proficiency in a foreign language. Unfortunately, there is no record of replications by Smith or others. Whether the magnitude of Smith's findings squelched any other initiatives can only be conjectured. It should be pointed out that manufacturers continued to include expansion and compression capabilities in the "deluxe" models of their equipment. It can only be concluded that many practitioners appreciated these features and purchased them, although they had no independent, empirical confirmation of their effectiveness.

28.8.2.4.1. Summary. Twelve studies conducted since the passage of the NDEA in 1958 were discussed in this section. They differed considerably in scale, populations, and methodology. Although all concerned language laboratories in some way, they did not all seek to answer the same questions other than the general one of effectiveness. For these reasons, it is difficult to draw a conclusion. This body of research does not offer clearcut confirmation of the utility of language laboratories, yet neither does it suggest that they are detrimental to language learning. Perhaps the inconclusiveness of the record is because the investigations that were conducted were not following an agreed-upon agenda. The larger educational technology community began the period with such an agenda (Allen, 1959; Meierhenry, 1962). This lack of focus was costly: Pederson (1987) claimed that it was the lack of solid research concerning courseware that led to the decline of language laboratories.

It would be hasty, however, to dismiss all language laboratory research. It can readily be determined that the use of audio resources within the foreign-language community has differed significantly from that of the larger educational technology community. Not surprisingly, this different use fostered different research. What was unique to the utilization and study of audio resources within foreign-language circles? One can first note the interest in psychophysics and the acoustic parameters of equipment. Besides Buka, Freeman, and Locke (1962) and Freeman and Buka (1965), who were discussed above, Hayes (1963) should be mentioned. He culled a wide range of human factors literature in order to offer standards to be used in laboratory purchase specifications. At this time, the broader educational technology community was more concerned with visual rather than auditory perception. A clear example of this pictorial bias is the fifth issue of volume 10 of the *Audio-Visual Communication Review* (1962), which was entitled "Perception Theory and AV Education." It contained no mention of the aural sense. Such a slanting of interest belies the "audio" component in the name of the flagship journal of the educational technology field at the time. Saettler's *The Evolution of American Educational Technology* shows that this inclination persists; visual media are accorded much more attention than are audio media.

Related to acoustic and perceptual matters are equipment features. Some of the studies reviewed in this section of the chapter (e.g., Young & Choquette, 1965; Chomei & Houlihan, 1970) were concerned with this issue. This is also unique to the body of language laboratory research. Only the studies of compressed and expanded speech showed an interest in machine capabilities.

The reader must be aware that in foreign-language teaching, audio materials are not used merely to present declarative information in an alternative medium; rather audio is the medium. What is meant is that language as normally used in everyday life is predominantly oral and aural rather than visual. Knowledge of a language is more properly considered to be procedural; such knowledge must be inculcated and learned by doing. This perhaps explains why the foreign-language community showed little, if any, interest in the channel effectiveness literature and audio-tutorial instruction (ATI). That the latter was not pursued is somewhat surprising when one considers that Smith (1970), in his essay in the prominent ACTFL yearbook, had proposed that ATI held promise for better foreign-language courseware.

The record shows that periodically foreign-language practitioners have been diligently involved in the use of audio technology. The record also shows that this enthusiasm has

waned, and the blame was put on the lack of adequate research. In order to avoid a repeat of this mistake, it would appear that research is required. Audio technology is still in use: less and less alone, but more and more in computer-based multimedia systems. These systems permit the easy collecting of usage data. One can hope that audio technology research will enter a new phase with this new equipment.

28.8.3 Conclusions about Language Labs

According to Last (1989), "language teachers as a body have been more ready than most to accept and explore the pedagogical potential of new technologies as they have emerged" (p. 15). Yet Richards and Nunan (1992) judge that "technology at present is underexploited in language learning and teaching" (p. 1203). It is not the purpose of this chapter to resolve this contradiction. Suffice it to say that there is evidence for both positions, and one has reason to be optimistic about the future. In 1990 the U.S. Department of Education funded the first National Foreign Language Resource Center. Two centers, the University of Hawaii and San Diego State University, have been offering workshops on the use of technology. The FLAME (Foreign Language Applications in Multimedia Environment) project at the University of Michigan is another reason for hope.

At the outset of this portion of the chapter, it was stated that the larger educational technology community has not fully appreciated the history of language laboratories. The scant attention paid to them in Saettler's *The Evolution of American Educational Technology* was cited to support this point. Nor has the history of the research that accompanied language laboratories been acknowledged heretofore. The proof of this contention can be seen in Allen's (1971) review of past research. This essay in the AVCR by its long-time editor contained no mention of what were then recent studies, some of which had attracted much attention in the popular press. It is hoped that this summary will contribute to setting the record straight.

REFERENCES

Adelson, L. (1975). Comprehension by college students of time compressed lectures. *Journal of Experimental Education 44,* 53–59.

Agency for International Development (n.d.). *Interactive radio instruction.* Newton, MA: Education Development Center.

Aikens, H.F. & Ross, A.J. (1977). Immediate, repetitive playback/record—a practical solution. *NAALD Journal 11* (2), 40–46.

Allen, E.D. (1960). The effects of the language laboratory on the development of skill in a foreign language. *Modern Language Journal 44,* 355–58.

Allen, W.H. (1959). Research on new educational media: summary and problems. *Audio-Visual Communication Review 7,* 83–96.

— (1971). Instructional Media research: past, present, and future. *Audio-Visual Communication Review 19,* 5–18.

Altamura, N.C. (1970). Laboratory a liability. *French Review 43,* 819–20.

Anderson, E.W. (1964). Review and criticism. *Modern Language Journal 48,* 197–206.

Bagster-Collins, E.W., et al. (1930). *Studies in modern language teaching.* New York: Macmillan.

Barrutia, R. (1967). The past, present, and future of language laboratories. *Hispania 50,* 888–99.

Bates, A.W. (1983). Adult learning from educational television: the open university experience. *In* M.J.A. Howe, ed. *Learning from television: psychological and educational research,* 213–27. Washington, DC: American Psychological Association.

Bauer, E.W. (1964). A study of the effectiveness of two language laboratory conditions in the teaching of second year German. *International Review of Applied Linguistics 2,* 99–112.

Bell, R. (1969). An analysis of certain elements of an audio-take approach to instruction. Unpublished doctoral dissertation, University of Washington.

Benathy, B.H. & Jordan, B. (1969). A classroom laboratory instructional system (CLIS). *Foreign Language Annals 2,* 466–73.

Benz, C.R. (1971). Effects of time compressed speech upon the comprehension of a visual oriented television lecture. Unpublished doctoral dissertation, Wayne State University.

Boyle, V.A. (1969). Visual stimulation and comprehension of compressed speech. (Doctoral dissertation, George Peabody College for Teachers, 1969). Dissertation Abstracts International, 30, 5221B.

Breed, P.A. (1977). The relative effect of the controlled reader and the speech compressor on reading rate and comprehension. Unpublished Doctoral Dissertation, Northern Illinois University.

Brushwood, J. & Polmantier, P. (1953). *The effectiveness of the audio-laboratory in elementary modern language courses.* Columbia, MO: The University of Missouri.

Buchanan, M.A. & MacPhee, E.D. (1928). *An annotated bibliography of modern language methodology.* Toronto, Canada: University of Toronto Press.

Buka, M., Freeman, M.K. & Locke, W.N. (1962). Language learning and frequency response. *International Journal of American Linguistics 28,* 62–79.

Button, G.E. (1991). Audio-tutorial biology, andragogy, and self-esteem: relationships among independent and dependent variables. Dissertation Abstracts International, 53/02, 457. (University Microfilms No. AAC 9211068.)

Carroll, J.B. (1963). Research on teaching foreign languages. *In* N.L. Gage, ed. *Handbook of research on teaching,* 1060–1100. Chicago, IL: Rand McNally.

— (1969a). Modern languages. *In* R.L. Ebel, ed. *Encyclopaedia of educational research, 4th ed.,* 866–78. New York: Macmillan.

— (1969b). What does the Pennsylvania foreign language research project tell us? *Foreign Language Annals 3,* 214–36.

Carter, K.R. & Cooney, J.B. (1983). The audio-tutorial applied in psychology classes to minimize individual differences and save class time. *Teaching of Psychology 10,* 201–04.

Carver, R.P. (1973). Understanding, information processing, and learning from prose materials. *Journal of Educational Psychology 64,* 76–84.

— (1982). Optimal rate of reading prose. *Reading Research Quarterly 18* (1), 56–88.
— (1973). Effect of increasing the rate of speech presentation upon comprehension. *Journal of Educational Psychology 65,* 118–26.
Chomei, T. & Houlihan, R. (1970). Comparative effectiveness of three language lab methods using a new equipment system. *AV Communication Review 18,* 160–68.
Chall, J.S. & Dial, H.E. (1948). Predicting listener understanding and interest in newscasts. *Educational Research Bulletin 27,* 141–53.
Cicardo, A.R. (1974). An analysis of retention of a compressed speech message among various intelligence levels, rates of compression, and retention levels. Unpublished doctoral dissertation, Northwestern State University of Louisiana.
Clark, R.E. (1983). Reconsidering research on learning from media. *Review of Educational Research 53,* 445–59.
Clarke, C.C. (1918). The phonograph in modern language teaching. *Modern Language Journal 3,* 116–22.
Claudel, C.A. (1968). The language laboratory. *In* J.S. Roucek, ed. *The study of foreign languages,* 219–36. New York: Philosophical Library.
Constantine, M. (1964). Radio in the elementary school. *Science Education 48,* 121–32.
Cook, D.C. & Nemzek, C. (1939). Effectiveness of teaching by radio. *The Journal of Educational Research 33* (2), 105–09.
Cook, H.R. (1964). The effects of learning of structural drills in Spanish broadcast via high frequency AM radio (NDEA Title VII Project No. 1018). Bloomington, IN: Indiana University.
Cutler, R.L., McKeachie, W.J. & McNeil, E.B. (1958). Teaching psychology by telephone. *The American Psychologist 13* (9), 551–52.
Davies, N.F. (1982). Foreign/second language education and technology in the future. *NAALD Journal 16* (3/4), 5–14.
Davison, W.F. (1973). *The language laboratory: a bibliography, 1950-1972.* Pittsburgh, PA: University Center for International Studies and The English Language Institute, University of Pittsburgh.
Derthick, L.G. (1959). The purpose and legislative history of the foreign language titles in the National Defense Education Act, 1958. *Publications of the Modern Language Association 74,* 48–51.
Diekhoff, J.S. (1965). *NDEA and modern foreign languages.* New York: Modern Language Association.
DeKieffer, R.E. (1973). The national center for audio tapes. *In* J.W. Brown, ed. *Educational Media Yearbook,* 56–59. New York: Bowker.
DeWick, H.N. (1935, Jun.). The relative recall effectiveness of visual and auditory presentation of advertising material. *The Journal of Applied Psychology 19,* 245–64.
Diehl, C.F., White, R.D. & Burke, K. (1959). Rate and comprehension. *Speech Monographs 26,* 229–32.
Dodge, J.W. (1968). Language laboratories. *In* E.M. Birkmaier, ed. *Britannica review of foreign language education, Vol. 1,* 331–35. Chicago, IL: Encyclopaedia Britannica.
Driscoll, J. (1981). Research trends in rate-controlled speech for language learning. *NALLD Journal 15* (2), 45–51.
Duker, S. (1974). *Time compressed speech.* Metuchen, NJ: Scarecrow.
Eckhardt, W.W. (1970). Learning in multi-media programmed instruction as a function of aptitude and instruction rate controlled by compressed speech. Unpublished doctoral dissertation, University of Southern California.
Eddy, F.D. (1944). The language studio. *Modern Language Journal 28,* 338–41.
Ek, J.D. (1974). Grant fever. *NALLD Journal 9* (1), 17–23.
Ely, P. (1984). *Bring the lab back to life.* Oxford, England: Pergamon.
Erickson, C.I. & King, I. (1917). A comparison of visual and oral presentations of lessons in the case of pupils from third to ninth grade, *School and Society 6* (136), 146–48.
Eshgh, R., Hoxeng, J., Provenzano, J. & Casals, B., eds. (1988). Radio-assisted community basic education (RADECO). Pittsburgh, PA: Duquesne University Press.
Fairbanks, G., Everitt, W.L. & Jaeger, R.P. (1954). Method for time or frequency compression-expansion of speech. *Transactions of the Institute of Radio Engineers, Professional Groups on Audio, AU 2,* 7–12.
—, Guttman, N. & Miron, M.S. (1957). Auditory comprehension of repeated high speech messages. *Journal of Speech and Hearing Disorders 22,* 23–32.
Fergen, G.K. (1954). Listening comprehension at controlled rates for children in grades iv, v, and vi. Unpublished doctoral dissertation, University of Missouri.
Fotos, J.T. (1955). The Purdue laboratory method in teaching beginning French classes. *Modern Language Journal 39,* 141–43.
Foulke, E. (1968). Listening comprehension as a function of word rate. *Journal of Communication 18,* 198–206.
—, ed. (1966). *Proceedings of the Louisville conference on time-compressed speech.* Louisville, KY: Center for Rate Controlled Recordings.
—, Amster, C.H., Bixler, R.H. & Nolan, C.Y. (1962). The comprehension of rapid speech by the blind. *Exceptional Children 11,* 134–41.
Foulke, E.A. & Sticht, T.G. (1967). *The intelligibility and comprehension of accelerated speech.* Proceedings of the Louisville Conference on Time Compressed Speech, Louisville, KY, 21–28.
— (1969). *Proceedings of the second Louisville conference on rate and/or frequency controlled speech.* Louisville, KY: University of Louisville. (ERIC Document Reproduction Service No. Ed 61 682.)
— (1962, Mar.). *A comparison of two methods of compressing speech.* Symposium at the Southeastern Psychological Association, Louisville, KY.
— & Sticht, T.G. (1969). A review of research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin 72,* 50–62.
Fox, B. & Roblin, J.S. (1952). Retention and sleep. *Journal of Experimental Psychology 43* (1), 75–79.
Freeman, M.Z. & Buka, M. (1965). Effect of frequency response on language learning. *AV Communication Review 13,* 289–95.
Friedman, H.L., Freedle, R.O., Norris, C.M. & Orr, D.B. (1966). *Further research on speeded speech as an educational medium.* (Report No. AIR-E-50-7-66-TR-3.) Silver Spring, MD: American Institute for Research in the Behavioral Sciences (ED 044 903).
—, Orr, D. & Norris, C. (1966). *Further research on speeded speech as an educational medium—the use of listening aids.* (Report No. 3.) Silver Spring, MD: American Institutes for Research in Behavioral Sciences (ED 044 903).
Funke, E. (1949). Rebuilding a practical phonetics laboratory.

German Quarterly 21, 120–25.
Furnham, A. & Gunter, B. (1985). Sex, presentation mode and memory for violent and non-violent news. *Journal of Educational Television 11,* 99–105.
— & Gunter, B. (1987). Effects of time of day and medium of presentation on recall of violent and non-violent news. *Applied Cognitive Psychology 1,* 255–62.
—, Benson, I. & Gunter, B. (1987). Memory for television commercials as a function of the channel of communication. *Social Behaviour 2,* 105–12.
—, Gunter, B. & Green, A. (1988). Remembering science: the recall of factual information as a function of presentation mode. *Applied Cognitive Psychology 4,* 203–12.
—, Proctor, T. & Gunter, B. (1988). Memory for material presented in the media: the superiority of written communication. *Psychological Reports 63,* 935–38.
Gallacher, T. & Stevens, R. (1954). Teaching with a tape recorder. *California Journal of Secondary Education 29,* 312–14.
Garvey, W.D. (1953). The intelligibility of speeded speech. *Journal of Experimental Psychology 45,* 102–08.
George, R.G. (1970). Retention of prose material as a function of rate of presentation and difficulty of material. *Audio Visual Communication Review 18,* 291–99.
Gibson, R. (1958). Tape recordings are used to teach seventh-grade students in Westside Junior-Senior High School, Omaha, NE. *Bulletin of the National Association of Secondary Principals 42,* 81–93.
— (1959). The tape recordings experiment is expanded in Westside Junior-Senior High School, Omaha, NE. *Bulletin of the National Association of Secondary Principals 43,* 49–72.
— (1960). Final report on the Westside High School Teaching-by-Tape Project. *Bulletin of the National Association of Secondary Principals 44,* 56–62.
Godfrey, E.P. (1967). *The state of audiovisual technology: 1961–1966.* Washington DC: Department of Audiovisual Instruction, National Education Association.
Goldhaber, G. (1967). Listener comprehension of compressed speech when the difficulty level of the content and the sex of the listener varied. Unpublished master's thesis, University of Maryland.
— (1970). Listener comprehension of compressed speech as a function of the academic grade level of the subjects. *The Journal of Communications 20,* 167–73.
Goldstein, H. (1940). Reading and listening comprehension at various controlled rates. *Teacher's College Contributions to Education,* No. 821.
Green, J.R. (1965). Language laboratory research: a critique. *Modern Language Journal 49,* 367–69.
Grittner, F. (1964). The shortcomings of language laboratory findings in the IAR-Research Bulletin. *Modern Language Journal 48,* 207–10.
Gullette, C.C. (1932). Ear training in the teaching of pronunciation. *Modern Language Journal 16,* 334–36.
Gunter, B. (1987). *Poor reception: misunderstanding and forgetting broadcast news.* Hillsdale, NJ: Erlbaum.
— & Furnham, A. (1986). Sex and personality differences in recall of violent and non-violent news from three presentational modalities. *Personality and Individual Differences 7,* 829–37.
—, — & Gietson, G. (1984). Memory for the news as function of the channel of communication. *Human Learning 3,* 265–71.
—, — & Leese, J. (1986). Memory for information from a party political broadcast as a function of channel of communication. *Social Behaviour 1,* 135–42.
Hartman, F.R. (1961). Single and multiple communication: a review of research and proposed model. *Audio Visual Communication Review 9,* 235–62.
Harvey, T.E. (1978). The matter with listening comprehension isn't the ear: hardware & software. *NALLD Journal 13* (1), 8–16.
Hayes, A.S. (1963). *Language laboratory facilities: technical guide for the selection, purchase, use, and maintenance.* Washington, DC: U.S. Department of Health, Education, and Welfare.
Heinich, R. (1968). The teacher in an instructional system. *In* F.G. Knirk & J.W. Childs, eds. *Instructional technology: a book of readings,* 45–60. New York: Holt.
Hirsch, R. (1954). *Audio-visual aids in language teaching.* Washington, DC: Georgetown University Press.
Hocking, E. (1964). *Language laboratory and language learning.* Washington, DC: Division of Audiovisual Instruction, National Education Association.
Holmes, G. (1985). From the president. *NALLD Journal 19* (2), 5–7.
Huppert, J. & Lazarowitz, R. (1990). Pupils abilities and academic achievement in an individualized biology curriculum. *Research in Science and Technological Education 8,* 117–31.
Hutchison, J.C. (1961). *Modern foreign languages in high school: the language laboratory.* Washington DC: U.S. Department of Health, Education, and Welfare.
Jarlett, F.G. (1971). The falsely accused language laboratory: 25 years of misuse. *NALLD Journal 5* (4), 27–34.
Johnston, M.C. & Seerley, C.C. (1960). *Foreign language laboratories in schools and colleges.* Washington, DC: U.S. Department of Health, Education, and Welfare.
Klavon, A.J. (1975). Time-compressed lecture: an alternative for increased teacher learner interaction. Unpublished doctoral dissertation, University of Maryland.
Keating, L.C. (1936). Modern inventions in the language program. *School and Society 44,* 677–79.
Keating, R.F. (1963). *A study of the effectiveness of language laboratories.* New York: Teachers College, Columbia University.
Kelly, L.G. (1969). *25 centuries of language teaching.* Rowley, MA: Newbury.
Kenner, R. (1981). Report on the Concordia Colloquium on language laboratories. *NALLD Journal 16* (2), 15–18.
Koekkoek, B.J. (1959). The advent of the language laboratory. *Modern Language Journal 43,* 4–5.
Kramer, E.J. & Lewis, T.R. (1951). Comparison of visual and non-visual listening. *Journal of Communication 1* (2), 16–20.
Kulik, J.A., Kulik, C.-L.C. & Cohen, P.A. (1979). Research on audio-tutorial instruction: a meta-analysis of comparative studies. *Review of Educational Research 11,* 321–41.
Lange, D.L. (1968). Methods. *In* E.M. Birkmaier, ed. *Britannica Review of Foreign Language Education, Vol. 1,* 281–310. Chicago, IL: Encyclopaedia Britannica.
Larsen, R.P. & Feder, D.D. (1940). Common and differential factors in reading and hearing comprehension. *The Journal of Educational Psychology 31,* 241–51.
Leshan, L. (1942). The breaking of a habit by suggestion during sleep. *The Journal of Abnormal and Social Psychology 37,* 406–08.
Last, R.W. (1989). *Artificial intelligence techniques in lan-*

guage learning. Chichester, England: Horwood.
LeMon, R.E. (1986). Computer labs and language labs: lessons to be learned. *Educational Technology 26,* 46–47.
Leon, P.R. (1962). *Laboratoire de langues et correction phonétique.* Paris: Didier.
Levin, L.M. (1931). More anent the phonetic laboratory method. *Modern Language Journal 15,* 427–31.
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.
— (1965). Comments on "language laboratory research: a critique." *Modern Language Journal 49,* 369–70.
Loder, J.E. (1937). A study of aural learning with and without the speaker present. *Journal of Experimental Education 6,* 46–60.
Ludrick, J.A. (1974). A study of the effects of controlled delivery instruction upon the achievement of college students using compressed speech audio and television pictorials. Unpublished doctoral dissertation, the University of Oklahoma.
Lysaght, C.E. (1969). Geriatrics: effect of speech rate and pacing procedures upon the responses to verbal stimuli by three age groups. Unpublished doctoral dissertation, Boston University.
Maclachlan, J. (1982). Listener perception of time-compressed spokespersons for radio commercials. *Journal of Advertising 22,* 47–51.
May, M.A. & Lumsdaine, A.A. (1958). *Learning from films.* New Haven, CT: Yale.
McCoy, I.H. & Weible, D.M. (1983). Foreign languages and the new media: the videodisc and the microcomputer. *In* C.J. James, ed. *Practical applications of research in foreign language teaching,* 105–52. Lincolnwood, IL: National Textbook.
Meierhenry, W.C. (1962). Needed research in the introduction and use of audiovisual materials: a special report. *Audio-Visual Communication Review 10,* 307–16.
Miller, G.A. & Lichlinder, J.C. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America 22,* 167–73.
Molstad, J. (1955). Readability formulas and film grade-placement. *Audio-Visual Communication Review 3,* 99–108.
Mosenthal, P. (1976–77). Psycholinguistic properties of aural and visual comprehension as determined by children's abilities to comprehend syllogisms. *Reading Research Quarterly 12,* 55–92.
Mueller, T. & Borglum, G. (1956). Language laboratory and target language. *French Review 29,* 322–31.
Menne, J.W., Klingenschmidt, J.E. & Nord, D.L. (1969, Mar.). *The feasibility of using taped lectures to replace class attendance.* Paper presented at the annual meeting of American Educational Research Association, Los Angeles.
Nasser, D.L. & McEwen, W.J. (1976). The impact of alternate media channels: recall and involvement with messages. *AV Communication Review 24,* 263–72.
Nelson, H.E. (1948). The effect of variations of rates on the recall by radio listeners of straight newscast. *Speech Monographs 1,* 173–80.
Newmark, M. (1948). Teaching materials: textbooks, audio-visual aids, the language laboratory. *In* M. Newmark, ed. *Twentieth century modern language teaching,* 456–62. New York: Philosophical Library.
Nichols, R.G. & Stevens, L.A. (1957). *Are you listening?* New York: McGraw-Hill.
Novak, J.D. & Musunda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal 28,* 117–53.
Nugent, G.C. (1982). Pictures, audio, and print: symbolic representation and effect on learning. *Educational Communication and Technology Journal 30,* 163–74.
NHK Radio-Television Cultural Research Institute (1956). *The listening effect of radio English classroom.* Tokyo: NHK.
Otto, S. (1989). The language laboratory in the computer age. *In* W.F. Smith, ed. *Modern technology in foreign language education: applications and projects,* 13–41. Chicago, IL: National Textbook.
Parker, W.R. (1960). Foreword. *In* F.J. Oinas, ed. *Language teaching today,* pp. v–viii. Bloomington, IN: Indiana University Research Center in Anthropology, Folklore, and Linguistics.
— (1961). *The national interest and foreign languages,* 3d ed. Washington, DC: U.S. Department of State.
Pederson, K.M. (1987). Research on CALL. *In* W.F. Smith, ed. *Modern media in foreign language education: theory and implementation,* 99–131. Chicago, IL: National Textbook.
Peebles, S. (1938). *The phonetics laboratory and its usefulness.* Unpublished MA thesis. Boulder, CO: University of Colorado.
Peterson, P. (1974). Origins of the language laboratory. *NALLD Journal 8* (4), 5–17.
Olsen, D.R. & Bruner, J.S. (1974). Learning through experience and learning through media. *In* D.R. Olsen, ed. *Media and symbols,* 125–50. Chicago, IL: University of Chicago Press.
Orr, D.B. & Friedman, H.L. (1964). *Research on speeded speech as an educational medium.* (Progress Report, Grant No. 7-48-7670-203.) U.S. Department of Health, Education, and Welfare, Office of Education, Washington, DC.
—, — & Williams, J.C. (1965). Trainability of listening comprehension of speeded discourse. *Journal of Educational Psychology 56,* 148–56.
Popham, W.J. (1962). Tape recorded lectures in the college classroom II. *AV Communication Review 10,* 94–101.
Porter, J.J. & Porter, S.F. (1964). A critique of the Keating report. *Modern Language Journal 48,* 195–97.
Racle, G.L. (1976). Laboratoire de langues: problèmes et orientations. *Canadian Modern Language Review 32,* 384–88.
Postlethwaite, S.N. (1970). The audio-tutorial system. *American Biology Teacher 32,* 31–33.
— (1972). The audio-tutorial system: incorporating minicourse and mastery. *Educational Technology 12* (9), 35–37.
— (1975). Students are a lot like people! *American Biology Teacher 37,* 205.
— (1978). Principles behind the audio-tutorial system. *NSPI Journal 17,* 3–4,18.
— (1980). Improvement of science teaching. *BioScience 30,* 601–04.
—, Novak, J. & Murray, H.T., Jr. (1972). *The audio-tutorial approach to learning,* 3d ed. Minneapolis, MN: Burgess.
Potter, M.C. (1990). Remembering. *In* D.N. Osherson & E.E. Smith, eds. *Thinking: an invitation to cognitive science,* 3–32. Cambridge, MA: MIT Press.
Rao, P.V. (1977). Telephone and instructional communication. *In* I.D.S. Pool, ed. *The social impact of the telephone,* 473–86. Cambridge, MA: MIT Press.
Reid, R.H. (1968). Comprehension of compressed speech as a function of difficulty of material. Unpublished doctoral dissertation, Florida State University.
Richards, J. C. & Nunan, D. (1992). Second language teaching

and learning. *In* M.C. Aikin, ed. *Encyclopaedia of educational research,* 6th ed., 1200–08. New York: Macmillan.

Robertson, E.M. (1977). The effects of different rates of recorded speech on the listening comprehension of adult remedial readers. Unpublished doctoral dissertation, University of Georgia.

Rohwer, W.D. & Harris, W.J. (1975). Media effect on prose learning in two populations of children. *Journal of Educational Psychology 67,* 651–57.

Ross, R.A. (1964). Look at listeners. *Elementary School Journal 64* (7), 369–72.

Rossiter, C.M., Jr. (1972). Sex of the speaker, sex of the listener, and listening comprehension. *The Journal of Communication 22,* 64–69.

— (1970). The effects of rate of presentation on listening test scores for recall of facts, recall of ideas, and generation of inferences. Unpublished doctoral dissertation, Ohio University.

Rulon, P.V. (1943a) A comparison of phonographic recordings with printed material in terms of knowledge, *The Harvard Educational Review 8,* 63–76.

— (1943b). A comparison of phonographic recordings with printed material in terms of knowledge gained through their use in a teaching unit, *The Harvard Educational Review 8,* 163–75.

— (1943c). A comparison of phonographic recordings with printed motivation to further study. *The Harvard Educational Review 8,* 246–55.

Saettler, P. (1990). *The evolution of American educational technology.* Englewood, CO: Libraries Unlimited.

Salomon, G. (1979). *Interaction of media, cognition, and learning.* San Francisco, CA: Jossey-Bass.

— (1984). Television is "easy" and print is "tough": the differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology 76,* 647–58.

Sanchez, J. (1959). Twenty years of modern language laboratory (an annotated bibliography). *Modern Language Journal 43,* 228–32.

Schenk, E.H. (1930). Practical difficulties in the use of the phonetics laboratory. *Modern Language Journal 15,* 30–32.

Scherer, G.A.C. & Wertheimer, M. (1964). *A psycholinguistic experiment in foreign-language teaching.* New York: McGraw-Hill.

Schramm, W. (1972). What the research says. *In* W. Schramm, ed. *Quality instructional television,* 44–79. Honolulu: HI: University Press of Hawaii.

Scott, R.J. (1965). Temporal effects in speech analysis and synthesis. Unpublished doctoral dissertation, University of Michigan.

Short, H.S. (1977). A comparison of variable time-compressed speech and normal rate speech based on time spent and performance in a course taught by self-instructional methods. *British Journal of Educational Technology 8,* 146–57.

Sisson, C.R. (1970). The effect of delayed comparison in the language laboratory on phoneme discrimination and pronunciation accuracy. *Language Learning 20,* 69–88.

Smith, P.D. (1969a). The Pennsylvania foreign language research project: teacher proficiency and class achievement in two modern languages. *Foreign Language Annals 3,* 194–207.

— (1969b). An assessment of three foreign language teaching strategies and three language laboratory systems. *The French Review 43,* 289–304.

— (1980). A study of the effect of "slowed speech" on listening comprehension of French. *NALLD Journal 14* (3/4), 9–13.

Smith, W.F. (1970). Language learning laboratory. *In* D.L. Lange, ed. *Britannica review of foreign language education, Vol. 2,* 191–237. Chicago, IL: Encyclopaedia Britannica, Inc.

— (1987). *Modern media in foreign language education: theory and implementation.* Chicago, IL: National Textbook.

— (1989). *Modern technology in foreign language education: applications and projects.* Chicago, IL: National Textbook.

Stack, E.M. (1964). The Keating report: a symposium. *Modern Language Journal 48,* 189–210.

— (1971). *The language laboratory and modern language teaching.* New York: Oxford University Press.

Stanton, F.N. (1934). Memory for advertising copy presented visually and orally, The *Journal of Applied Psychology 18,* 45–64.

Stich, T.G. (1968). Some relationships of mental aptitude, reading ability, and listening ability using normal and time-compressed speech. *Journal of Communication 18,* 243–58.

— & Glasnap, D.R. (1972). Effects of speech rate, selection difficulty, association strength, and mental aptitude on learning by listening. *The Journal of Communication 22,* 174–88.

— (1971). Failure to increase learning using the time saved by the time compression of speech. *Journal of Educational Psychology 62,* 55–59.

Stiefel, W.A. (1952). Bricks with straw—the language laboratory. *Modern Language Journal 36,* 68–73.

Strei, G. (1977). Reviving the language lab. *TESOL Newsletter 11,* 10.

Svoboda, R.G. (1978). Audio-tutorial courses for college algebra and trigonometry: a progress report. Fort Wayne, IN: Indiana University. (ED 167125.)

Taylor, S.E. (1965). Eye movements in reading: facts and fallacies. *American Educational Research Journal 2,* 187–202.

Tripp, S.D. (1994, Aug. 5). Do media affect memory? Paper presented at the 3d Practical Aspects of Memory Conference, College Park, MD.

Valette, R.M. (1969). The Pennsylvania project, its conclusions and its implications. *Modern Language Journal 53,* 396–404.

Voor, J.B. & Miller, J.M. (1965). The effect of practice on the comprehension of time compressed speech. *Speech Monographs 32,* 452–54.

Waltz, R.H. (1930). The laboratory as an aid to modern language teaching. *Modern Language Journal 15,* 27–29.

— (1931). Language laboratory administration. *Modern Language Journal 16,* 217–27.

— (1932). Some results of laboratory training. *Modern Language Journal 16,* 299–305.

Wasserman, H.M. & Tedford, W.H. (1973). Recall of temporally compressed auditory and visual information. *Psychological Reports 32,* 499–502.

Watts, M.W. Jr. (1971). Differences in educational level and subject matter difficulty in the use of compressed speech with adult military students. *Adult Educational Journal 21,* 27–36.

Whitehouse, R.S. (1945). The workshop: a language laboratory. *Hispania 28,* 88–90.

Williams, D.L., Moore, D. M., & Sewell, E.H., Jr. (1983–84). Effects of compressed speech on comprehension of community college students. *Journal of Educational Technology Systems 12,* 273–84.

Wolcott, L.L. (1993). Audio tools for distance education. *In* B. Willis, ed. *Distance education: strategies and tools,* 135–64, Washington, DC: American Psychological Association.

Wood, C.D. (1966). Comprehension of compressed speech by elementary school children. Indiana University, Bloomington, Indiana (ED 003 216.)

Worchester, D.A. (Jan., 1925). Memory by visual and auditory presentation. *The Journal of Educational Psychology 16* (1), 18–27.

Young, W.E. (1936). The relation of reading comprehension and retention to hearing comprehension and retention. *Journal of Experimental Education 5* (1), 30–39.

Young, C.W. & Choquette, C.A. (1965). An experimental study of the effectiveness of four systems of equipment for self-monitoring in teaching French pronunciation. *International Review of Applied Linguistics* 3, 13–49.