DOCUMENT RESUME

ED 039 515 AL 002 402

AUTHOR	Pulliam, Robert
TITLE	Automatic Speech Recognition in the Teaching of Second Languages: An Annotated Bibliography.
INSTITUTION	Pulliam and Associates, Research Consultants, Fairfax, Va.
REPORT NO	Monograph-2
PUB DATE	Jan 70
NOTE	27p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.45
DESCRIPTORS	Acoustic Phonetics, *Annotated Bibliographies, *Computer Assisted Instruction, *Language Instruction, Language Laboratories, Psycholinguistics, Teaching Machines, Teaching Methods
IDENTIFIERS	*Automatic Speech Recognition

ABSTRACT

This bibliography, containing 73 annotated entries, is designed as an aid to researchers who may wish to experiment using automatic speech recognition in programed instruction, computer assisted instruction, or task simulation devices. The bibliography samples recent literature on the technology of automatic speech recognition, on efforts to employ it at elementary technical levels, and on devices which evaluate speech as visual displays. Collateral material is added, from several disciplines, which may be useful to experimenters in formulating further research. A special focus is the teaching of second languages. A brief introduction deals with the development of automatic speech recognition and its use in teaching machines. (Author/FWB)

PULLIAM AND ASSOCIATES / RESEARCH CONSULTANTS

10242 STRATFORD AVENUE • FAIRFAX, VIRGINIA 22030 • (703) 591-7645

AUTOMATIC SPEECH RECOGNITION IN THE TEACHING OF SECOND LANGUAGES: AN ANNOTATED BIBLIOGRAPHY

Robert Pulliam

Monograph Number Two - January 1970

CONTENTS

RESEARCH AND PROFESSIONAL REPRESENTATION IN WASHINGTON, D.C. AND VICINITY
LANGUAGE SCIENCE • EDUCATIONAL TECHNOLOGY • INFORMATION SYSTEMS

TABLE OF CONTENTS

INTRODUCTION

Mechanical recognition of human speech has been discussed hypothetically for many years, as a necessary feature of any ideal automatic teaching system. This bibliography is designed as an aid to researchers who may wish to experiment using automatic speech recognition in programmed instruction, computer assisted instruction, or task simulation devices.

When Homer Dudley demonstrated the first Vocoder in 1928, it was clear that automatic speech recognition was theoretically possible, and would be achieved with time. It was not until 1952, however, that K.H. Davis exhibited a working digit recognizer, and not until quite recently that practical, economic, and reliable equipment has been seen. From about 1966, such devices began to appear with increasing frequency. In the 1970's they will provide direct input of human instructions to an increasing variety of computer and control devices. Already automatic speech recognition has been used in astronaut maneuvering systems, in a device for requesting stock market quotations by telephone, and in sorting packages in the post office. One of its most valuable potential uses is a means of communication between student and machine in advanced educational systems.

Speech is the basic mode of human communication and symbolic behavior. With its correlate, aural comprehension, it preceeds writing and other graphic communication historically, and in individual human development. Speech is the most frequently used means by which humans describe reality, manipulate it in symbols, and seek to influence the behavior of others. It is of course the most used tool of classroom instruction.

It is therefore a recognized weakness of teaching machines that they do not yet provide for the student to speak. While they can present, as output, a variety of aural and visual stimuli; they recognize as input only rigidly structured manipulations (mostly keyboard) which must themselves be learned before other learning can commence. Lack of a spoken input has prevented realization of ideal student-machine transactions, especially for teaching the young, the handicapped, and students of language skills.

This bibliography samples recent literature on the technology of automatic speech recognition, on efforts to employ it at elementary technical levels, and on devices which evaluate speech as visual displays. Collateral material is added, from several disciplines, which may be useful to experimenters in formulating further research. A special focus of the bibliography is the teaching of second languages.

PART I

AUTOMATIC SPEECH RECOGNITION

A. Current Development Programs

Known programs directed toward realization of hardware for the automatic recognition of speech are listed:

1. Anke, D., and Hoeschele, P. "Simple Recognition Devices for the Spoken Numbers Zero to Nine." Kybernetik, 4 (June, 1968) 228-234.

   An effort in Germany, using German language input. With 24 filter channels, Anke claims 54-95% reliability, depending on the speaker.

2. Bobrow, Daniel G., Hartley, Alice K., and Klatt, Dennis H. A Limited Speech Recognition System II, NASA Contract NAS 12-138, Report No. 1819, 1 April 1969. Cambridge, Mass.: Bolt, Beranek and Newman, Inc., 1969

   One of two efforts by Bolt, Beranek and Newman Inc., using a device termed the LISPER. In its final configuration the system uses a modified pattern recognition technique, to identify up to one hundred words or brief phrases, and signal the phrase heard. Accuracy approaches 97%. Speaker dependence remains a problem.

3. Bobrow, Daniel G., and Klatt, Dennis H. A Limited Speech Recognition System, NASA Contract NAS 12-138, Report No. 1667, 15 May 1968. Cambridge, Mass.: Bolt, Beranek, and Newman, Inc., 1968.

   This report describes Bolt, Beranek and Newman work prior to May 1968.

4. Bobrow, Daniel G., and Klatt, Dennis H. "A Limited Speech Recognition System." Proceedings of the Fall Joint Computer Conference, 1968, 305-317. Washington, D.C.: Thompson Book Co., 1968.

   A general summary of Bolt, Beranek and Newman work, in a publicly available source.

5. Glenn, James W. Automatic Speech Recognition, A State of the Art Survey. Reston, Virginia: SCOPE Inc., October, 1969. (Unpublished MSS.)

An unpublished manuscript containing a concise survey of work on limited-segment recognition.

6. Hill, F.J., McRae, L.P., and McClellan, R.P. "Speech Recognition as a Function of Channel Capacity in a Discrete Set of Channels." Journal of the Acoustic Society of America, 44 (July, 1968) 13-18.

An experimental system developed at the University of Arizona, Tucson, using a tactile signal as output. Designed to permit cutaneous recognition by the deaf.

7. Information Sciences Laboratory Staff, SCOPE Inc. Automatic Speech Interpretation. Reston, Va.: SCOPE, Inc., 1968.

This sales document overstates the SCOPE capability, which is nevertheless as good as any in the field. The SCOPE device is pattern recognition based. It is programmed by a "training sequence," in which actual speech samples are provided as input, and pattern data is derived without the intervention of a human programmer. Recognition of up to 64 utterances is claimed, with a very economical cost in memory and processing. Some speaker independence has been realized.

8. Kopstein, Felix F. "Computers and Instruction at HumRRO." Educational Technology, 9 (July, 1969) 8.

The Human Resources Research Office at Alexandria, Virginia, has been developing a speech recognition program for use with the CAI effort described by Kopstein. No written report of the HumRRO speech work exists.

9. Lindgren, Nilo. "Machine Recognition of Human Language: Part I - Automatic Speech Recognition." Institute of Electronic and Electrical Engineers Spectrum, (March, 1965) 114-136.

Lindgren is the best general summary, both of speech recognition work and of the basic research in several disciplines upon which that work depends.

10. Oppenheim, A.V. "Speech Analysis-Synthesis System Based on Homomorphic Filtering." Journal of the Acoustic Society of America, 45 (February, 1969) 458-465.
This development at the Lincoln Laboratories, (MIT) is a digital-based study, not currently attempting a recognition device, and implying use in communications coding. Findings would clearly apply as input to a computer program for the digital recognition of single utterances.

11. Pulliam, Robert. Application of the SCOPE Speech Interpreter in Experimental Educational Programs. Pulliam & Associates Monograph No. 1. Fairfax, Va.: Pulliam & Associates, 1969.
A non-technical description of the SCOPE capability, addressed to the education market.

12. Reddy, D.R. "Computer Recognition of Connected Speech." Journal of the Acoustic Society of America, 42 (August, 1967) 329-347.
Recognizing the difficulty of identifying an acoustically satisfactory phoneme, this research seeks to reclassify phoneme groups in mathematically meaningful sets. A phoneme-based system would presumably lead to recognition of connected speech in terms of an orthographic transcript.

13. Sakai, T., et al. "Fundamental Studies of Speech Analysis and Synthesis." American Annals of the Deaf, 113 (March, 1968) 156-167.

14. Shoup, June E. Personal Communication. December 22, 1969.
The Speech Communications Laboratory, Inc. at Santa Barbara, California, is working toward speech recognition systems. No findings have been published, but descriptions of the work and anticipated hardware realizations are expected to be published by others in March, 1969.

15. Strong, William J. "Machine-aided Formant Determination for Speech Synthesis." Journal of the Acoustic Society of America, 41 (June, 1967) 1434-1443.

Speech synthesis from coded data is the obverse of speech recognition. A capability to decode automatically yields a capability to encode, and a capability to encode can probably yield a decode program. Strong's work at the U.S. Air Force Cambridge Research Laboratories, Bedford, Massachusetts, uses a human operator in a semi-automatic system, dependent on formant analysis. It recognizes phonemes at rates of 96% accuracy for vowels and 83% for consonants.

16. Teacher, C.F., Kellett, H., and Focht, L. "Experimental Limited Vocabulary Speech Recognizer." Institute of Electronic and Electrical Engineers: Audio and Electroacoustics, AU-15 (September, 1967) 127-130.

Teacher describes developments at Philco-Ford Corporation. It is understood that this project has been discontinued.

17. Terhardt, E. "A Contribution to the Automatic Detection of Spoken Numerals." Kybernetik, 3 (September, 1966) 136-143.

This project in Germany uses a function model of the human ear for detection of digits spoken in German. The ear model was built by Zwicker. Results, using a 24 channel filter input, were highly speaker dependent.

B. Limited Objective Experiments

At least two experiments have been conducted in which an interface device reacted to student speech, at a level less than that of identification of utterances. They are of interest in demonstrating the importance of a spoken response, and as guidance in the design of experiments with more sophisticated technology.

18. Buiten, Roger, and Lane, Harlan. "A Self-Instructional Device for Conditioning Accurate Prosody." International Review of Applied Linguistics, III (1965) 205-219.

Work by Harlan Lane and others at the Center for Research on Language and Learning Behavior of the University of Michigan is especially interesting. The SAID system developed there comes more closely than any other known work to an educational application of speech recognition. This quality-of-utterance experiment studied retention of phonemic behavior, learned under various strategies, after the subject had returned to his native language environment.

The SAID equipment measured phoneme formation in three dimensions: (1) Average speech power. (2) Frequency discrimination by a two-filter system. (3) Temporal spacing recognized by timed switches. The student's pronunciation was recorded on tape; acceptability of each parameter was displayed, one feature at a time, on a zero-center meter, and the student was invited to "shape" his phoneme formation in a series of trys.

19. Garvey, Catherine J., Johansen, Patricia A., and Noblitt, James S. A Report of the Developmental Testing of a Self-Instructional French Program. Washington, D.C.: Center for Applied Linguistics, October, 1967.

This early report of the French Self-Instructional Language Project at the Center for Applied Linguistics details (among other things) the procedure by which ideal achievable program frame strategies were devised, and a device modified for presenting the program. The materials are unique in the care which was taken to plan psychological strategies and linguistic sequence. Avoiding, insofar as possible, the dilemma of letting a machine dictate the kind of interaction which can take place between program and student, the researchers finally secured modification of the Appleton-Century-Crofts "Portable Laboratory System". Significant was the fact that the researchers recognized the necessity of spoken response as the primary behavior mode. The device was therefore fitted with a microphone, output of which triggered a voice operated relay. In effect it simply recognized that a subject did speak, and signalled for the next frame when he finished speaking. Thus when the program prompted a student to speak, program logic stopped other operations until an utterance had been recognized as having been attempted. Of course the device could make no judgement of the accuracy of the utterance, and would advance the program to the next frame even if the student said something irrelevant. (See further annotation of this same publication at 46).

20. Johansen, Patricia A. "The Development and Field Testing of a Self-Instructional French Program." The Linguistic Reporter, Supplement 24 (December, 1969) 13-27. Washington, D.C.: The Center for Applied Linguistics, 1969.

A current report, and best general reference on the French Self-Instructional Materials for the Center for Applied Linguistics (See Garvey above, 19). The interface is not described in detail.

C. Visual Display Systems

Visible display devices must be examined in any study of speech recognition, because historically they have been essential tools for the study of speech sound, and because they offer alternative means of evaluating the speech of students. Display devices can be categorized roughly as: (1) Recorders, which make a running record of acoustic events. (2) Transient displays, typically on cathode ray tubes, which can either display features continuously as they occur, or can isolate and display momentarily a single feature or segment.

21. Barton, George W., and Barton, Stephen H. "Forms of Sounds as Shown on an Oscilloscope by Roulette Figures." Science, 142 (1965) 1455-1456.

Barton & Barton formed "roulette figures", which are specialized lissajous figures, formed by input to the X and Y axes of a cathode ray tube from a simple phase-shifting RC bridge. Figures are circular or oval for pure tones, but take on distinctive forms due to signal components other than the dominant formant. Variations can be seen for dialect and personal differences. The patterns formed were termed "caligraphones".

22. Cohen, Martin L. "The ADL Sustained Phoneme Analyzer." American Annals of the Deaf, 113 (March, 1968) 247-252.

The ADL Sustained Phoneme Analyzer is developed by Arthur D. Little, Inc., of Cambridge, Massachusetts, as a device for training the deaf using a cathode ray tube display.

23. Jensen, Paul G., and Westermeir, Fraz X. The Effect of Visual Feedback on Pronunciation in FL Learning, Project Termination Report, Macalester College, St. Paul, Minn. St. Paul, Minn.: Macalester College, 1968.

In a study at Macalester College, St. Paul, Minnesota, the Barton & Barton system was to examine visible feedback in the teaching of German. The project was terminated when it was felt that patterns generated did not correlate adequately with speech as perceived. Visual patterns were similar for sounds significantly different to the ear, and sounds considered similar on linguistic criteria made sharply different visual patterns. Furthermore, the visual patterns lost their identity much sooner

than the aural image, and seemed psychologically inferior to perceived speech as a means of feedback and evaluation. Visual display of the speech spectrum is suggested as an alternative, or visual feedback based on recognition rather than direct display.

24. Koenig, W., Dunn, H.K., and Lacy, L.Y. "The Sound Spectrograph." Journal of the Acoustic Society of America, 17 (1946) 19-49.

The sound spectrograph, as originally proposed by R.K. Potter, was demonstrated in 1946, and has since that time been the primary research tool in the study of speech sounds. It records frequency on the vertical scale, with intensity of the trace indicating energy, and time on the horizontal scale. It is interesting to note, especially in the early studies, how such sounds as bird songs form clear and distinctive patterns, but human speech forms patterns which are not clearly patterned, and which require painstaking study to decode.

25. Liberman, A.M., et al. "Why are Speech Spectrograms Hard to Read?" American Annals of the Deaf, 113 (March, 1968) 127-133.

This work was performed at Haskins Laboratories (New York), and at the University of Connecticut. It demonstrates that the difficulty of "phonemic" displays is predictable from and parallel to that of reading voice spectrograms, and is due to the low ratio of significant acoustic clue data to the volume of total information of the speech signal. It is concluded that the best visual displays of speech might be representations of the articulatory muscle contractions.

26. Lindgren, Nilo. "Machine Recognition of Human Languages: Part I. Automatic Speech Recognition." Institute of Electronic and Electrical Engineers Spectrum, (March, 1965) 114-136.

This article contains a quick historical treatment of the work which has been accomplished with the speech spectrograph.

27. Phillips, N.D., et al. "Teaching of Information to the Deaf by Visual Pattern Matching." American Annals of the Deaf, 113 (March, 1968) 239-254.

Three studies at Northwestern University (Chicago) find that any visual pattern display device for the deaf should be: (1) Displayed by log of frequency, since linear differences are less significant in higher ranges. (2) Should have a memory, or trace-holding feature, for display of the student performance. (3) Should have a storage feature for holding of model data. (4) Should display only voiced portions of the speech signal (!). (5) Should display model and response data together in the same display. (6) Should have means for normalizing the display for age, sex and individual differences.

28. Pickett, J.M., and Constrom, A. "A Visual Speech Trainer with Simplified Indication of Vowel Spectrum." American Annals of the Deaf, 113 (March, 1968) 120.

The Gallaudet Visual Speech Trainer (Gallaudet College, Washington, D.C.) displays speech in a trainer, using parameters of voice pitch, timing patterns, and certain distinctions from spectral distribution.

29. Pronovost, Wilbert. The Development and Evaluation of Procedures for Using the Voice Visualizer as an Aid in Teaching Speech to the Deaf, Final Report, Contract OEG-1-6-062017-1588. Boston, Mass.: University of Massachusetts, School of Education, 1967.

The Pronovost device is a roulette pattern type indicator developed by Lerner, and used in tandem with an audio signal. It was useful in the teaching of vowels, voiced and voiceless fricatives. A more encouraging study than others (such as Jensen, 23) it suggests that roulette figure feedback is useful only for certain selected features, and interferes with learning when not required.

30. Pronovost, Wilbert, et al. "The Voice Visualizer." American Annals of the Deaf, 113 (March, 1968) 230-238.

A summary of Pronovost's work published in a readily available source.

31. Risberg, Arne. "Visual Aids in Speech." American Annals of the Deaf, 113 (March, 1968) 178.

Visual aids developed at the Royal Institute of Technology in Stockholm are described, which teach sounds by successive approximations. Several types of indicator are mentioned. (1) A bar-graph of the frequency spectrum by energy level is derived at one point in time, and held in a temporary display by a memory device. (2) A fricative indicator is a high-frequency skewed, simplified bar-graph indicator, without a memory function. (3) An S indicator has been available commercially in Sweden for over ten years, and indicates an acceptably pronounced S sound with a yes/no indicator, variable for threshold level. (4) An Intonation indicator. (5) A rhythm indicator.

PART II

COLLATERAL REFERENCES

A. History

Research and development prior to 1965 is primarily of historical interest, but useful in understanding the difficulties which speech recognition presents, and the reasons for current directions in the art.

32. Davis, K.H., Biddulph, R., and Balashek, S. "Automatic Recognition of Spoken Digits." Journal of the Acoustic Society of America, 24 (1952) 637-642.

K.H. Davis gave the first public demonstration of a limited vocabulary speech recognizer at the June Conference on Speech Analysis at Massachusetts Institute of Technology in 1952. It recognized digits zero through nine.

33. De Lattre, Pierre. "Research Techniques for Phonetic Comparison of Languages." International Review of Applied Linguistics, I/2 (1963) 85-97.

A researcher, who has been a pioneer in the field for two decades, surveys acoustic research to 1963.

34. Golden, Roger M. "Vocoder Filter Design: Practical Considerations." Journal of the Acoustic Society of America, 43 (April, 1968) 803-810.

Cites work as early as 1928.

35. Le Histe, Ilse. Readings in Acoustic Phonetics. Cambridge, Mass.: MIT Press, 1967.

A book-length collection of significant published work in acoustic phonetics, containing 32 articles which systematically review work in acoustic phonetics between 1946 and 1960.

36. Lindgren, Nilo. "Machine Recognition of Human Languages: Part I. Automatic Speech Recognition." (Previously cited, 26).

First of a series of three excellent articles, a highly readable summary of work in several disciplines to 1965. Recommended as a first reference for persons new to the field.

B. Research Suggesting Need for Speech Recognition

37. Glaser, Robert W., Ramage, William W., and Lipson, Joseph I. The Interface Between Student and Subject Matter. Pittsburg, Ohio: Learning Research and Development Center, University of Pittsburg, 1966.

This is the best available study of the student/machine interface. Glaser studies the modality of present and future mechanical environments, and makes recommendations concerning the need to limit dependence on keyboards, add auditory entering behavior, and specifically asks for a speech recognition capability in advanced interface devices.

38. Glaser, Robert W., and Ramage, William W. "The Student Machine Interface in Instruction." Institute of Electronic and Electrical Engineers International Convention Record, Part 10. New York: Institute of Electronic and Electrical Engineers, 1967.

A brief statement, parallel to the above monograph, in a generally available source.

39. Peterson, Gordon E. "On the Nature of Speech Science." Annual Bulletin, 1967, Research Institute of Logopedics and Phoniatrics. Tokyo, Japan: Faculty of Medicine of the University of Tokyo, 1967.

Peterson speaks of the interdisciplinary nature of speech science, classifies its parts, and shows the relation to and dependence on electronic engineering and information science. He stresses the importance of developing vocoder systems and visual speech displays, and the general implications of "speech automation" with its two constituents: automatic synthesis, and automatic recognition of speech. Together, these capabilities will provide the means of communication with a computer, using natural speech.

40. Suppes, Patrick. Computer-Assisted Instruction in the Schools: Potentialities, Problems and Prospects, Technical Report No. 81, October 29, 1965. Stanford, Calif.: Stanford University Institute for Mathematical Studies in the Social Sciences, 1965.

Suppes identifies a problem of "stimulus deprivation" in most programmed and other machine instruction. He sees that foreign languages have already achieved machine instruction with the language laboratory, but there they have no individualization of instruction, and no evaluation of the overt response. He comments on the inadequacy of devices for presenting sound, and implies need for its counterpart in sound recognition.

41. Teslaar, A.P. van. "Learning New Sound Systems: Problems and Prospects." International Review of Applied Linguistics in Language Teaching, III/2 (1965) 79-93.

In an outstanding basic reference on sound, for the language teacher or other researcher not familiar with the disciplines concerned. Van Teslaar summarizes the field, noting that some 2% of voice sound is significant signal. In closing he makes a strong case against the language laboratory as conventionally used, and suggests it must at least be "audio active" and give prominence to features which are unique, contrastive, or likely to be incorrectly formed as a result of native language conditioned perception. He notes the inability of typical speakers to perceive their own distortions of second languages, and suggests Harlan Lane's SAID experiment as an alternative direction.

C. Selected CAI and PI in Foreign Languages

Listed are experiments in Computer Assisted Instructions (CAI) and Programmed Instruction (PI) in foreign language, which have implications for advanced techniques using automatic speech recognition.

42. Adams, E.N., Morrison, H.W., and Reddy, J.M. "Conversation with a Computer as a Technique of Language Instruction." Modern Language Journal, 52 (January, 1968) 3-15.

Pedagogical assumptions of the Adams technique are discussed. In this and subsequent entries by Adams, Morrison, and Rosenbaum (43, 44, 49). A CAI approach originally developed in German at the IBM Watson Research Center (Yorktown Heights, N.Y.) is treated. The technique used a teletype keyboard and CRT display, and was essentially a reading and writing laboratory, although limited use of spoken stimulus from a tape recorder was attempted. A subsequent experiment, based on the Adams approach with modifications by Adams and Rosenbaum, was run in Russian at The Defense Language Institute (West Coast Branch) at Monterey, California. That experiment, not yet reported in the literature, was interesting because of the high weekly student exposure to CAI and because a large body of materials (language software) was quickly compiled by language teachers, who had no prior understanding of computers and no training in a programming language.

Dr. Adams characterized his approach as "Conversation with a computer." The routines he devised offer an efficient cueing system, economic of machine effort, and credible as simulation of an interpersonnal exchange. But as demonstrated at the State University of New York at Stoneybrook, and at the Monterey School of Languages, they must be criticized precisely because they do not achieve "conversation" in an acceptable psychological or linguistic sense. Interchanges are reading and writing, which move too slowly to be psychologically comparable to conversation, and do not involve articulatory and sensory activity normal to spoken language - precisely the behavior to be taught.

In studying this and the next two references, researchers can speculate as to what might have been accomplished if Adams and Rosenbaum could have incorporated a more reliable speech output device, and a limited speech recognition feature, into their system.

43. Adams, E.N. The Use of CAI in Foreign Language Instruction, IBM Research Paper No. RC 2377. Yorktown Heights, N.Y.: Thomas J. Watson IBM Research Center, October 30, 1968.

Technical design of the German program. See Adams, item 42 above.

44. Adams, E.N. "Field Evaluation of the German CAI Laboratory." In Computer Assisted Instruction: A Book of Readings. New York: Academic Press, 1969.

Evaluation of the German field experiment. See Adams, item 42 above.

45. Atkinson, Richard C., and Suppes, Patrick. Program in Computer Assisted Instruction, Final Report USOE Contract No. OEC-4-6-061493-2089. Stanford, Cal.: Stanford University, August, 1968.

This final report to the U.S. Office of Education on the basic Suppes experiment at Stanford does not cover the related experiment by Joseph Van Campen in Russian (See Van Campen, entry 52), but describes the machine environment on which that experiment was run.

46. Garvey, Catherine, J., Johansen, Patricia A., and Noblitt, James S. A Report of the Developmental Testing of a Self-Instructional French Program. (Previously cited, item 19).

The Garvey-Johansen Self-Instructional French Program, developed at the Center for Applied Linguistics, is possibly the most credible and carefully evaluated set of programmed instruction (PI) materials in a foreign language, and could provide a software base for either a CAI or PI experiment using speech recognition, since the design includes regular spoken responses.

47. Johansen, Patricia A. "The Development and Field Testing of a Self-Instructional French Program." (Previously cited, item 20).
A summary report of the Garvey/Johansen work. See Garvey, item 46 above.

48. Morrison. H.W., and Adams, E.N. "Pilot Study of a CAI Laboratory in German." Modern Language Journal, 52 (May, 1968) 279-287.
Discussion of the Adams-Rosenbaum work, oriented toward language educators. See Adams, item 42.

49. Rosenbaum, Peter S. "The Computer as a Learning Environment For FL Instruction." Foreign Language Annals, 2 (May, 1969) 457-465.
Rosenbaum's comments on CAI in languages and most recent publication in the line of experiment concerned. See Adams, item 42.

50. Suppes, Patrick, and Jerman, Max. "Computer Assisted Instruction at Stanford." Education Technology, 9 (June, 1969) 22-24.
See Atkinson and Suppes, item 45.

51. Suppes, Patrick and Moringstar, Mona. "Computer Assisted Instruction." Science, 166 (17 October 1969) 343-350.
Data on results of the Stanford experiment in Russian. See Atkinson and Suppes, item 45.

52. Van Campen, Joseph. Project for the Application of Mathematical Learning Theory to Second Language Aquisition with Particular Reference to Russian, Project Report, USOE Contract No. C-0-8-00120901806. Stanford, Cal.: Stanford University, 1967.
Design of the Stanford Russian experiment. See Atkinson and Suppes, item 45.

D. The Language Laboratory

By 1963 the Language Laboratory had become a standard feature of modern second language teaching programs. Doubts about its effectiveness, at least as sometimes used, were raised first by the Keating Report and then by the Pennsylvania studies. In 1969 that issue became probably the principle subject of discussion in the profession. Automatic speech recognition is seen as a principle means of ensuring that the language laboratory can provide the effective individual practice for which it was designed.

53. Clark, John L.D. "The Pennsylvania Study and the Audio-Lingual versus Traditional Question." Modern Language Journal, 53 (October, 1969) 388-396.

The Modern Language Journal for October 1969 featured a series of commentaries on the Pennsylvania studies and the future of the Language Laboratory. This and three others (54, 56 and 59 below) are cited here.

54. Hocking, Elton. "The Laboratory in Perspective: Teachers, Strategies, Outcomes." Modern Language Journal, 53 (October, 1969) 404-410.

Comments on the Pennsylvania Studies (See Clark, above 53).

55. Keating, Raymond F. A Study of the Effectiveness of Language Laboratories. Columbia University, N.Y.: Institute of Administrative Research, Teachers College, 1963.

The text of the Keating Report. It suggests (by implication) ways in which a more responsive laboratory might be attempted in experiments using speech recognition.

56. Lange, D.L. "Methods." In The Britannica Review of Foreign Language Education, Vol. I. Chicago: American Council on the Teaching of Foreign Languages and Encyclopedia Britannica, Inc., 1968.

Discusses the Language Laboratory in its relation to methods in language teaching, with comments on the Keating and Pennsylvania Studies.

57. Smith, Phillip D., Jr., and Baranyi, Helmut A. A Comparison Study of the Effectiveness of the Traditional and Audio-Lingual Approaches to Foreign Language Instruction, Using Laboratory Equipment, Final Report, USOE Project No. 7-0133. Washington, D.C.: Educational Resources Information Center, 1968.

With Smith and Berger following, this constitutes the full text of the Pennsylvania Studies.

58. Smith, Phillip D., Jr., and Berger, Emanuel. An Assessment of Three Foreign Language Learning Strategies, Using Three Language Laboratory Systems, Final Report, USOE Project No. 5-0683. Washington, D.C.: Educational Resources Information Center, 1968.

59. Vallette, Rebecca M. "The Pennsylvania Project, Its Conclusions and Its Implications." Modern Language Journal, 53 (October, 1969) 396-404.

Comments on the Pennsylvania Study by a leading expert in language testing. (See Clark, item 53).

E. Acoustic, Phonetic, Neurophysiological Research

60. House, Arthur S., and Fairbanks, Grant. "The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels." Journal of the Acoustic Society of America, 25 (1953) 105-113.

Illustrative of early work implying overlap of phoneme categories in vowels. Both consonant environment and speaker differences affect the locus of formants.

61. Mundie, J. Ryland, and Moore, Thomas J. Speech Analysis as the Ear Sees It: Aural Topography, Paper delivered to the 1967 Conference on Speech Communication and Processing of the Institute of Electronic and Electrical Engineers. Cambridge, Mass.: Massachusetts Institute of Technology, 1967.

Mundie worked with the cochlear analog, and is one of several researchers who see physiological and neurological transforms of speech data .s useful in resolving signal cues from otherwise non-decodable sound data.

62. Mundie, J. Ryland. Personal Communication. December 16, 1969.

Mundie uses the cochlear analog made by Dr. Stewart at Santa Rita Technology, Inc., Santa Clara, California.

63. Rabiner, Lawrence R. "A Digital-Formant Synthesizer for Speech Synthesis Studies." Journal of the Acoustic Society of America, 43 (April 1968) 822-828.

Work on synthesis-by-rule at Bell Telephone Laboratories, Murray Hill, New Jersey.

64. Rabiner, Lawrence R., Levitt, H., and Rosenberg, G. "Investigation of Stress Patterns for Speech Synthesis by Rule." Journal of the Acoustic Society of America, 45 (January, 1969) 92-101.

Attempts by Bell Telephone Laboratories to avoid "machine-like quality" of synthetic speech, by application of prosodic rules.

65. Shoup, Jean E. Allophones of Midwestern English. Vol. II of Automatic Speech Recognition, University of Michigan Engineering Summer Conferences, 1963, 529-531. Ann Arbor, Mich.: University of Michigan Press, 1963.

A more recent work, by a researcher involved in speech recognition studies, casting doubt on the validity of the phoneme except as a perceptual phenomenon.

66. Scott, Robert J. "Time Adjustment in Speech Synthesis." Journal of the Acoustic Society of America, 41 (January, 1967) 60-65.

Compares methods of compression or expansion of speech to normalize for speech and cadence, including the problem of frequency distortion when speech is expanded by slowed reproduction.

67. Stevens, Kenneth N., and Klatt, Mary M. Study of Acoustic Properties of Speech Sounds, Bolt, Beranek and Newman, Inc., Scientific Report No. 8, 30 August 1968. Cambridge, Mass.: Bolt, Beranek and Newman, Inc., 1968.

A systematic appreciation of English speech acoustics, sponsored by the Advanced Research Projects Agency, Department of Defense, specifically for application in speech recognition.

68. Stewart, J.L. Speech Processing with a Cochlear-Neural Analog, United States Air Force Aerospace Medical Lab Report 1-140, February, 1967. Wright-Patterson AFB, Ohio: 1967.

This earlier work on the cochlear analog, was used by Mundie, item 61.

69. Yilmaz, Huseyin. "A Theory of Speech Perception." Bulletin of Mathematical Biophysics, 29 (December, 1967) 793-825.

Yilmaz theorizes that speech perception must have developed in an evolutionary manner, analogous to biological evolution, and in response to the physical properties of sound, the distribution of its energy in the environment, and the tendency of sensory capability to optimize for survival. Conclusions are implied as to the general nature of speech perception. Yilmaz does not satisfactorily treat the speech process as a communications requirement, involving sound production simultaneously with perception, as the condition of survival.

70. Yilmaz, Huseyin. "A Theory of Speech Perception. II. Bulletin of Mathematical Biophysics, 30 (September, 1968) 455-479.

Further to Yilmaz, item 69 above.

F. Toward Connected Speech

All existing speech recognition hardware, and all theoretically attainable devices, deal with isolated segments of speech. The question of processing continuous speech, even to isolate specific fragments, is a larger problem not likely to be solved in a hurry. At its largest focus, the ability to process continuous random speech would require automata so nearly human in their capability as to be frightening. Some research in progress suggests a capability to deal with limited families of anticipatable sentences, in a manner useful for instructional machines.

71. Chapin, Paul G., and Norton, Lewis M. A Procedure for Morphological Analysis. La Jolla, Cal.: Published jointly by the University of California at La Jolla and the Mitre Corporation, July, 1968.

Deals with the analysis of English morphology by mathematically applicable rule.

72. Lindsay, Robert K. A Heuristic Parsing Tree for a Language Learning Program, Information Processing Report No. 12, University of Texas. Austin, Texas: University of Texas, May 28, 1964.

A program to simulate the language learning behavior of humans, using a "labelled dependency tree."

73. Von Glaserfield, Ernst, and Pisani, Pier Paolo. The Multistore System, MP-2, Georgia Institute for Research, Progress Report, U.S. Air Force Office of Scientific Research Grant No. 1319-67, November, 1968. Athens, Ga.: Georgia Institute of Research, 1968.

A parsing system is demonstrated. The system describes all English sentences in terms of a parsing tree with a "significant address" system of programming in which the location of bytes in core is significant in terms of English syntax, with resulting economy in accessing and processing. Program runs on an IBM 360/65 at the University of Georgis.