

LISTENING COMPREHENSION IN THE NATIVE LANGUAGE:
AN EXPERIMENTAL STUDY

by Michiyo Jean Yamashita

Thesis presented to the Department of
Linguistics and Modern Languages of
the University of Ottawa as partial
fulfillment of the requirements for
the degree of Master of Arts

Ottawa, May, 1973



UMI Number: EC55229

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EC55229
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

This thesis was written under the direction and guidance of Professor Fernand Marty. I wish to thank him and the members of my committee, Professors Th. R. Hofmann and P.C. Hauptman for their comments and criticisms, as well as support during moments of crisis. I also thank all the other members of the department who gave me timely advice and suggestions.

TABLE OF CONTENTS

- I -- Introduction
 - 1. Motivation of this Experiment
 - 2. Background Research in Hearing and Speech Perception
 - 2.1 Hearing and the Ear
 - 2.2 Speech Perception
 - 2.3 Masking
 - 3. Review of the Literature
 - 3.1 Experimentation
 - 3.2 Relation to Past Experiments
 - 3.3 Education

- II -- Experimental Design
 - 1. Subjects
 - 2. Test Instruments

- III -- Tabulation and Analysis of Data
 - 1. Listening Comprehension Test Analysis
 - 2. Auditory Discrimination Test Analysis
 - 3. Hypothesis Formulation Test Analysis
 - 4. Variables 17 - 19
 - 5. Statistical Analysis

- IV -- Findings
 - 1. Listening Comprehension Test
 - 2. Hypothesis Formulation Test
 - 3. Correlations

V -- Conclusions

1. General Conclusions
2. Implications of the Research
3. Suggestions for Further Research

Bibliography

Appendices

1. Forms for tests
2. Test answers
3. Correlations
4. Scattergrams

CHAPTER I

INTRODUCTION

1. MOTIVATION OF THIS EXPERIMENT

Listening¹ is a complex and active process. Its complexity can be seen in the fact that its factors or components have still to be thoroughly analyzed and its teaching is still a matter of debate among teachers and others in the education field. Thus tests of listening comprehension vary widely in content and objectives. (See Burger, 1972, for a review of tests together with some suggested techniques for testing.) Listening is active, that is, the listener plays a role, in that homophones (two words with different meanings but the same sound) are understood and recognized according to the listener's interpretation of the whole utterance. The context at the beginning of the sentence may impose one interpretation, and the end may suddenly switch interpretations. A classic example of this is the sentence "Rapid righting with his uninjured hand saved from loss the contents of the capsized canoe." (from Lashley, quoted in Thorpe, 1961, p. 5). When hearing the oral rendition, up to the phrase "saved from loss", people tend to understand "writing" not "righting".

The following line of investigation originated from Professor Marty's doubts as to the utility of films for language teaching if the objective of the course was to be understanding a face-to-face conversation with a native speaker. He found that listening comprehension in some students did not improve with exposure to films, even when

each film was repeated to near-memorization. Applied linguists seem to believe that recorded material in general is useful for students in any given language course. W.F. Mackey says, "People have been known to succeed in learning a second language and develop an acceptable accent in it, simply through the radio....Regular attendance at foreign films may succeed in maintaining a person's comprehension of the second language." (Mackey, 1965, p. 115). The people Mackey is talking about must be the exceptional, highly-motivated students who would learn regardless of or despite the method used. The language in films is quite different from both ordinary conversational and written language. It is usually much more emotional or dramatic and utterances cannot be repeated upon request. Success in listening comprehension of this type of language may require different abilities than for other forms of language. It was noted that all the listening comprehension tests which were examined tested comprehension of an orally-read text.

The present experiment tested the ability of native speakers to comprehend utterances under normal conditions of speaking and listening. The results of this test were analysed and correlated with subsequent tests of two factors hypothesized to be the basic components of listening comprehension: auditory discrimination and hypothesis formulation. We were interested in seeing whether or not these two factors could account for ability in listening comprehension. Such questions as the following

were also examined: Which was the more important of the two? Do the two factors complement each other in any way? That is, does an individual compensate for a deficiency in one by a superior ability in the other?

The experiment was performed for European French. The subjects were adult native speakers of French who had been away from France for an average of two and a half years. They were given three tests, a general listening comprehension test, an auditory discrimination test and a hypothesis formulation test. Materials for the Listening Comprehension Test were recorded from radio, television and movie sound tracks produced in France. The Auditory Discrimination Test tested discrimination of tones and synthetic vowels, and phonetic identification of short segments of speech. The Hypothesis Formulation Test consisted of clearly recorded sentences requiring unusual divisions of sound into words and of printed sentences requiring another possible meaning for the same sounds. Scores from all the sub-tests of these three tests were correlated with each other and also with personal data about the subjects themselves.

2. BACKGROUND RESEARCH IN HEARING AND SPEECH PERCEPTION

2.1 HEARING AND THE EAR

Helmholtz and early physicists were experimenting and theorizing on how the ear perceived sound as early as the nineteenth century (Helmholtz 1912). This line of research has continued into the present with the work of von Békésy (1957). Although there have not been decisive results to choose between the two major theories (the Place Theory and the Volley Theory), much knowledge has been gained about the anatomy of the ear, and more recently, about the functional operation of the ear. There is increasing evidence that all the major theories are correct in limited domains, and the focus of research is shifting to the means by which accurate judgements of pitch, loudness and duration are possible in a short time.

2.2 SPEECH PERCEPTION

Research into the perception of speech sounds in particular can be said to have begun with the invention of the sound spectrograph. Although there was considerable speculation before this based on fragmentary, limited or impressionistic data, not much serious work was possible until this invention made evidence readily available about the input to the ear.

Several other devices have proved valuable: the Pattern-Playback, which is a sound spectrograph in reverse; and more recently, computer-driven speech synthesizers and

analysers² (Denes & Pincus, 1963, pp. 125-147). This research has been directed at discovering how the ear extracts from a speech wave the elements that are distinctive in a given language.

One of the first beliefs to fall with the advent of plentiful and accurate data was that phonemes could be extracted in isolation; another was that there were pauses which demarcate phrases. A lot of other minor beliefs were destroyed about the specific nature of many phonemes, but the above two were important because it meant that no phoneme could be recognized without knowing what its neighboring phonemes were and that there was no way for the listener to identify how to cut up the speech train into words. Research went to seeking cues for recognizing distinctive features. These procedures for recognition under ideal conditions, are not better than 90% accurate.

This led inevitably to suspicion that the units the listener recognized were larger (Ladefoged, 1967), perhaps even phrase-sized units, and much effort in the psycholinguistics of speech perception has been devoted to discovering what is the size of the perceptual unit; distinctive feature, phoneme, word, phrase or sentence.

2.3 MASKING

Much research has been carried out on listening comprehension by specialists in psychology, "communication science", education and engineering. In fact, most of the

earliest experimental work was done by electrical engineers to aid in the design of communications equipment (radios and telephones). They were not interested in the listening comprehension ability of individuals, but in the quality of the channel and the intelligibility of the message under various conditions (Licklider & Miller, 1951).

It was the engineers who pioneered the technique of "masking" speech, that is, making speech less intelligible either by putting it through a filter to remove the high or low frequency components of energy (called "high-pass" and "low-pass" filtering, respectively), or by superimposing tones, other voices, or "white noise" (random noise, equally distributed in all frequencies) on the message (Miller, 1947, Licklider & Miller, 1951).

There followed many experiments which used this technique of masked or degraded speech; these were directed at determining how language perception works. Experimenters tried to determine the intelligibility of various types of material under different signal-to-noise (S/N) ratios. Miller et al, (1951) used numbers, words, sentences and nonsense syllables; Hirsch et al, (1954) used nonsense words, monosyllabic, disyllabic and polysyllabic words and sentences. The intelligibility of a word was measured by its resistance to noise -- i.e. a word that was intelligible at a S/N ratio of 5 db is more intelligible than another word which is intelligible only at a S/N ratio of 10 db. Experiments have determined the intelligibility of a word

based on its length (Rubenstein et al, 1959), its probability of occurrence (Rubenstein & Pollack, 1963), its context (Miller et al, 1951, Bruce, 1956). Also examined have been the effect of anticipation (Bruce, 1958), the effect of the rate of the utterance (Pickett & Pollack, 1963), and the word-frequency effect (Savin, 1962, Goldiamond & Hawkins, 1958, Brown & Rubenstein, 1961, Rosenzweig & Postman, 1958).

Conclusions: The general conclusions from these studies were that words were more easily recognized if they were known to come from a certain finite list -- the words being more intelligible if the number of possible choices were fewer--(Pollack et al, 1958, 1959), and that words in context (such as sentences) were more intelligible than words without context. Nonsense syllables were found to be the hardest type of material to understand. Longer words were also more intelligible because there are fewer alternatives for, say, a 5-syllable word, and also because each syllable provides a context for all the other syllables.

Another conclusion was that the word-frequency effect played an important part in speech recognition. When subjects' guesses for ambiguous stimuli were analysed, it was found that the guesses are usually frequent words. Thus the frequency of a given word (i.e. its probability of occurrence) influences the ease, speed and correctness with which it is identified.

3. REVIEW OF THE LITERATURE

3.1 EXPERIMENTATION

Masked speech was used in the Miller & Nicely experiment (1955) which established that English consonants show differing resistances to noise, and the study done by Spelsky (1968) in which he developed a listening comprehension test in English as a second language by using speech masked by white noise at S/N ratios from 50 db (no noise) to 1 db. He compared the performances of native and non-native listeners and found that this test was a good measure of listening proficiency in non-natives. His preliminary test sentences contained "phonological tricks", such as,

"Police washed the windows".

(Spelsky, 1968, p. 87) which may be heard under some noisy conditions as,

"Please wash the windows".

He later eliminated these as he claimed that they also were a form of masking.

Other research has demonstrated a heavy dependency of word-recognition on semantic context (Treisman, 1965) and has shown that semantic and syntactic context can completely override phonetic cues (Warren & Warren, 1970). Not only can a person correctly identify a word with a phoneme missing (replaced on the tape by a cough), but he also insists that every phoneme was heard,

i.e. he cannot even detect that it is missing. Further experimentation by Warren & Warren has shown that this holds true even for a whole syllable. "Phonemic restoration" occurs with a low-frequency word which has no higher frequency alternatives and which is found in a context leading to the expectation of that word. The semantic and grammatical context provided by a whole sentence affects the perception of a given word in that sentence (Miller & Isard, 1963). Lieberman (1963) showed that this was true for production as well as perception. Words taken from proverbs such as "A stich in time saves nine" tended to be pronounced less clearly and were therefore less intelligible than the same words in other sentences, such as "The number that you will hear is nine". At least one theory (Hofmann, 1971) claims to account for the use of lexical and grammatical redundancy in speech perception.

3.2 RELATION TO EARLIER EXPERIMENTS

Because there is no mention to the contrary, we assume that the test material in all these previous experiments (except Pickett & Pollack, 1963) was read in the normal reading style, viz. a moderately slow rate without the rapidity and the hesitations of normal speech. Our experiment used utterances that one would normally hear in a cinema, on television or on the radio. That is, they were not contrived for the purposes of a listening comprehension test. However, this immediately posed problems. There was a great deal of masking³ in these test utterances nonetheless. The material was from poorly recorded commercial products, and there was background noise, phonological tricks and a lack of context. However, there was no way to systematically measure or vary the masking present in each utterance, with the result that the percentage of subjects giving a correct answer varied from almost 0% on some items to 100% on others.

Perhaps the greatest form of "masking" in our test material was the removal of considerable context. This "context of situation" (Ladefoged, 1959) comprises preceding events and the preceding linguistic utterance(s), the situation of the listener and the speaker, the time and place of the speech act, and the expectation by the listener of something that will make sense. In the utterances used in this experiment, only the preceding events and linguistic acts were significant context. Even so, some of our

utterances, taken out of context as they were, really could not be said to make sense. The other facet of context, that is, the limitation of the possible phonetic and syntactic combinations by the structure of the language, was left intact.

3.3 EDUCATION

The main studies in education have been concerned with correlating scores from standardized listening comprehension tests, such as the Brown-Carlson Listening Comprehension Test, or specially-constructed tests, with scores in school subjects (Still, 1955), intelligence, and other individual qualities of the subjects (Nichols, 1948, Ainsworth & High, 1954, Blewett, 1951). Researchers have also tried to see whether or not listening comprehension as measured by these tests improves with classroom training (Ericksen, 1954, Johnson, 1951). Various studies conclude that the correlation of intelligence with listening comprehension is between .4 and .8, so that any correlation listening comprehension has with reading comprehension may be due to the common factor of intelligence which correlates with both of them (p. 12, Spearritt, 1962). Hearing ability as measured by pitch discrimination, loudness, rhythm, time, quality, auditory memory span and speech sound discrimination tests (Seashore and Templin) did not seem to correlate significantly with listening comprehension ability (Ainsworth & High, 1954). A good educational background

seems to be an important factor for good listening comprehension (p. 178, Wilkinson & Atkinson, 1966) although generally, it has been found that people really do understand very little of what they do hear -- college students about one-half of their lectures (Brown, 1950, Nichols, 1949) or less (Irvin, 1953). Cartier (1953) found in over 100 grade 10 students that only one-quarter of the students understood two-thirds of the material he presented, one-half understood about one-half the material, and the other quarter understood only about one-third of the material. He found no significant differences between males and females although at least one of these experimenters (Caffrey, 1955) found males superior to females in listening comprehension.

In his study, Nichols (1948) concluded that the main factors influencing listening comprehension ability of college freshmen for a lecture were; intelligence, ability to recognize correct English usage, size of vocabulary, ability to make inferences and to see the structure of a speech, interest, emotional adjustment to speaker's thesis and curiosity about the subject, and physical fatigue. The sex of the listener was listed under factors which might be connected to listening comprehension, along with high-school scholastic achievement, speech training, and susceptibility to distraction. Hearing acuity and previous training in the subject matter were not related to listening comprehension. In the conclusion, Nichols also notes that listening

comprehension can be improved by training, and that for the average untrained listener, slightly more than two-thirds of the material presented was understood.

Karlin (1942) gave a battery of hearing tests to try and find the auditory factor or factors for pitch, loudness, complexity and duration of sounds. However, instead of four factors, he isolated nine among the thirty-four scores.

Using the technique of factor analysis, Spearritt (1962) tested 400 boys and girls to see whether performance in listening comprehension could be simply a reflection of other abilities such as reasoning, verbal comprehension, attention and memory. He gave a series of tests for listening comprehension and the other abilities (here he used tests based on Karlin's work) together with tests of school achievement and hearing. He rejected his hypothesis as the analysis extracted a separate listening comprehension factor closely related to verbal comprehension and also related to span memory and rote memory, inductive reasoning and auditory resistance.

Notes to Chapter I

- ¹ "Listening" is understood to include hearing, identifying (these can be called speech perception), understanding and interpreting spoken language. It is also known as listening comprehension or "auding" (Caffrey, 1955).
- ² The most promising to-date is an analysis/synthesis programme for a PDP-11 made by The Speech Research Group of Carleton University.
- ³ We are using the term "masking" in Spolsky's sense, viz. a condition which hinders perception of a signal which might otherwise be clearly comprehended.

CHAPTER II

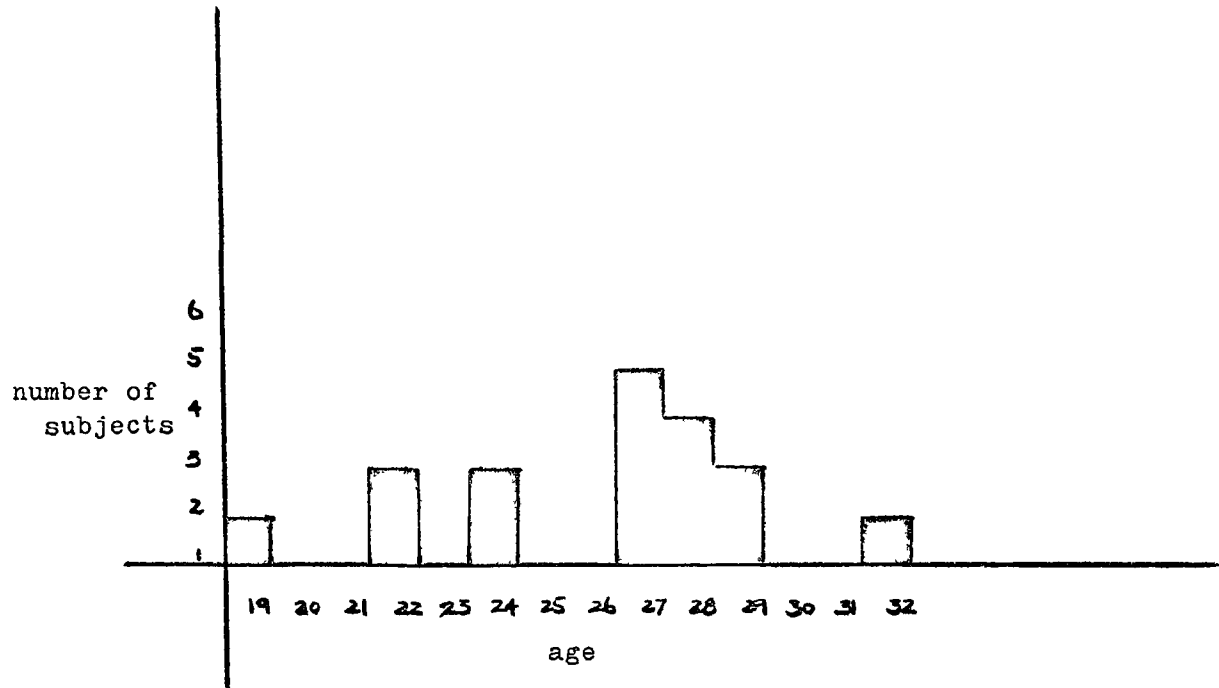
EXPERIMENTAL DESIGN

EXPERIMENTAL DESIGN

1. SUBJECTS

The subjects who participated in the experiment were native speakers of French who had been living out of France for no more than four years, and most of them had returned to France for visits during their holidays. Since all the available people who had these qualifications were asked to participate, this was an incidental sample. We tried to stay within a ten-year range in age--actually, the range was from 19 to 32 years, with the average age of the subjects being 26.2 years, the average deviation 2.5 years, the standard deviation 3.36 years. Table 1, Subject Age, shows a near-normal distribution of ages. All the subjects were from the university population and well-educated, the majority being teachers of French. Seventeen people in all took the first Listening Comprehension Test (in groups, when it was possible) in April 1972; however, only fifteen took the second test which was administered one to two months later, so that complete test results were available for these fifteen subjects, eight female, seven male. Furthermore, one of these subjects (S.12) was eliminated from the later analyses as testing conditions for this subject were not comparable to the others, so that complete data was compiled for fourteen subjects (8 female, and 6 male).

Table 1 - Subject Age



2. TEST INSTRUMENTS

All the test material was presented on tape. Recording and playback were done on professional quality equipment (Viking tape recorder, Ampex amplifier), with fidelity equal to or above NAFTB standards; in other words, very little distortion was produced by the equipment itself. Instructions were explained orally by the experimenter and subjects could ask questions at any time during the testing sessions: the Listening Comprehension Test took about one hour to complete; the Auditory Discrimination and Hypothesis Formulation Tests together lasted another hour.

2.1 The Listening Comprehension Test (Variables 1 - 6)

Listening comprehension tests usually consist of passages followed by test questions. These passages are well-prepared and well-enunciated in slow, clear language, without poorly-constructed sentences or sentence fragments. The questions may test not only comprehension but also memory for content, especially in longer passages of fifteen minutes or more. We were interested in the comprehension of films, where the type, speed, and clarity of language differs from the language of the above tests. Therefore, the Listening Comprehension Test was constructed from soundtracks of radio, television and films.

As each utterance differed from the next, thus eliminating any aid in comprehension from the context of situation, questions on the utterances could not be easily

made to test comprehension. Therefore, the technique of "shadowing" used, for instance, in the Miller & Isard (1963) experiment, was considered. A kind of modified shadowing was used to allow the test to be given in groups where the subjects could not repeat what they had heard. This modification consisted of having the subjects write down the utterance immediately after hearing it. Three tries were given to see if subsequent re-hearing would result in a significant improvement in comprehension.

The Listening Comprehension Test was made up of forty-five utterances taken from commercial disc recordings, movie sound tracks, radio and television programs selected for the difficulty we expected them to present in the absence of context or visual cues to help the listener understand the meaning. (See Appendix 2A for a complete list of the questions.) Some contained distorting or disrupting background noise; others presented unusual combinations of words. Each item was recorded three times. Subjects were asked to write down what they had understood after each playing, taking as much time as needed. Space was provided after each question for writing down comments, critiques, or guesses (e.g. number of syllables heard if the word was incomprehensible). Each question was marked as to completely right or wrong for each of the three trials, to produce scores for Variables one to four. Each question was again marked (impressionistically) on a 5-point scale; 0,

25, 50, 75, 100% correct - for each of the three trials to serve as data for Variables five and six. (See below.)

2.2 The Auditory Discrimination Test (Variables 7 - 10)

This test measured auditory discrimination in various ways in order to discover which facets of auditory discrimination, if any, might correlate with listening comprehension scores. In order to test for auditory discrimination, it was decided to test for hearing acuity by using pure (sinusoidal) tones, for ability to identify vowels by presenting vowels in isolation, and for ability to handle short speech sounds by presenting segments of real speech. The first section (having three parts), presented recorded pairs of pure tones (Variable seven). This format provided a test independent of language ability. It tested ability to hear various frequencies and the differences between them. Subjects were asked to indicate whether the second tone of each pair was higher or lower, and louder or softer than the first tone. Part A presented four pairs of tones, each being played twice; Part B presented the same four pairs, but in a different order and played only once. Part C presented two series of rising, falling and level tones. For each series, subjects were asked to indicate which of four schema presented on the answer sheet represented the melody of the tones.

The second section tested ability to make phonemic recognitions of vowels in not too perfect

renditions of language-like material. The term "not too perfect" is used because the sounds were synthesized on a speech synthesizer. The speech synthesizer was used to eliminate the possibility of mistakes in production or pitch changes that a human voice might make. This section (Variable eight) consisted of synthesized French vowels. Subjects were given a list of eight words, exemplifying the vowels /i y u E e ø o a/, and were asked to write down the word containing the vowel they heard. Each vowel was played twice in succession. In all these sections, questions were marked either wrong or right: scoring was based on the number of correct answers.

The third section of the auditory discrimination test (Variables nine and ten) tested accuracy of linguistic discriminations. It consisted of short segments of sound taken from the original Listening Comprehension Test. These segments ranged in length from three to seven phonemes. Because the segments were short, there was not enough context to allow the subject to use the redundancies inherent in sentence structure, syntax, lexicon and phonological structure. Even the experimenter found it difficult to identify these segments in isolation, although the segments were perfectly clear in the complete utterance. Subjects were asked to write down what they heard in any orthography they wished to use. Each segment was played twice. This section was corrected phoneme by phoneme, the

total score being the total number of phonemes correctly identified.

2.3 Hypothesis Formulation Test (Variables 11 - 16)

This was designed to test the ability to discover different ways to associate a given sound with meaning. It consisted of three sections, one testing ability to anticipate and the other two ability to formulate hypotheses. The first section consisted of five passages taken from plays, where the last word or words was cut out. Subjects were asked to complete the last sentence of each passage. This section was later eliminated (See Hypothesis Formulation Test Analysis, Ch. III, section 3.)

Nichols (1948) concluded that listening comprehension was related to ability to make inferences. We therefore designed the following sections to test this conclusion. In the second section, (Variables 11 - 13) sixteen sentences were read on tape in a clear voice and at a moderately slow speed by a native French speaker. (The reading rate was about 3.7 syllables per second.) These sentences were composed of unusual syntactic combinations or unexpected lexical combinations. As examples, the syntactic pattern NP+NP+V, NP+V is found in "Qui terre a, guerre a". An unexpected lexical combination is found in "Ce gout doux est pis que tout". Discovering the correct answer depended mainly on the subjects' abilities to make the proper syntactic divisions and think of other possible words

fitting a given combination of sounds. For example, hearing /fɔ̃/, one often thinks of "font", but there is also the verb form "fond". For each question subjects were given twelve seconds to write down what had been said, then were later given a chance to rehear any question they wished to hear, with one minute allowed for each question. Sixteen questions were recorded, but one (Q7, see Appendix 2C) was eliminated from the scoring, because of the mispronunciation of one word.

This section was marked in three ways: first of all, the number of completely correct answers was tabulated with no credit given to partially correct answers. (Variable 11) Secondly, each question was graded on the 5-point scale (same as the one in the Listening Comprehension Test) and the scores summed across questions. Thirdly, the number of answers corresponding in every respect to the answer wanted was tabulated. For example, for Q15, "Les en louer m'est dur", five subjects wrote "Les ans loués mais durs", which was marked as correct on the first and second markings, but wrong on the third marking.

The third section consisted of ten printed sentences which have two or more possible interpretations when pronounced aloud. They were taken from examples given by Delattre (1966) and supplemented by examples collected by F. Marty. Subjects were asked to write down an interpretation other than the one given. For example, if the sentence was "C'est un ceuf", the subject could write

"C'est un neuf". Two minutes were allowed for completion of this section, with subjects being asked to try every question quickly and to come back to those which proved difficult. This section was marked as to the number of questions correctly answered out of ten.

CHAPTER III

TABULATION & ANALYSIS
OF
DATA

1. LISTENING COMPREHENSION TEST ANALYSIS (Variables 1 - 6)

The raw data (scores) for the listening comprehension test were organized according to subject and question in the following ways:

a) Table 2

This table shows the number of subjects who gave wrong responses for each of the trials of the 45 questions for the listening comprehension test. It shows that only questions Q14, Q38 and Q40 (indicated by (N) on the table) proved difficult for all the subjects, and even for these questions, there were subjects who managed to understand them. (S6 answered Q14 and S7 answered Q40 correctly). Several questions were understood by all the subjects by the third trials: Questions 3, 9, 11, 13, 15, 25, 37, 43 (indicated by (A) on the table).

Table 2 Listening Comprehension Test: Number of Subjects Giving Incorrect Answers for Each Question

Question	Trial 1	Trial 2	Trial 3	3 Trials
Q1	6	4	4	4
Q2	11	7	7	6
Q3 (A)	2	1	0	0
Q4	4	4	5	3
Q5	1	1	1	1
Q6	9	4	4	4
Q7	3	3	3	2
Q8	3	1	1	1
Q9 (A)	0	0	0	0
Q10	8	9	10	7
Q11 (A)	0	0	0	0
Q12	1	1	1	1
Q13 (A)	2	1	0	0
Q14 (N)	14	14	13	13
Q15 (A)	1	0	0	0
Q16	12	13	11	9
Q17	4	4	5	4
Q18	6	3	1	1
Q19	1	1	1	1
Q20	6	4	6	4
Q21	9	4	4	4
Q22	9	4	4	4
Q23	11	11	10	9
Q24	2	3	3	1
Q25 (A)	2	0	0	0
Q26	3	3	1	1
Q27 (A)	0	0	0	0
Q28	4	2	2	2
Q29 (A)	0	0	0	0
Q30	8	7	7	7
Q31	11	7	9	4
Q32	7	6	6	5
Q33	8	6	5	5
Q34	12	7	6	5
Q35	9	4	2	2
Q36	12	9	7	7
Q37 (A)	2	0	0	0
Q38 (N)	14	14	14	14
Q39	5	2	2	2
Q40 (N)	14	14	13	13
Q41	3	2	1	1
Q42	13	10	9	9
Q43 (A)	0	0	0	0
Q44	14	12	11	11
Q45	<u>6</u>	<u>3</u>	<u>4</u>	<u>3</u>
Totals:	272	203	193	170
Percentage of questions wrong:	43.2%		30.6%	
Difference between Trials 1 and 3:			12.5%	

b) Table 3

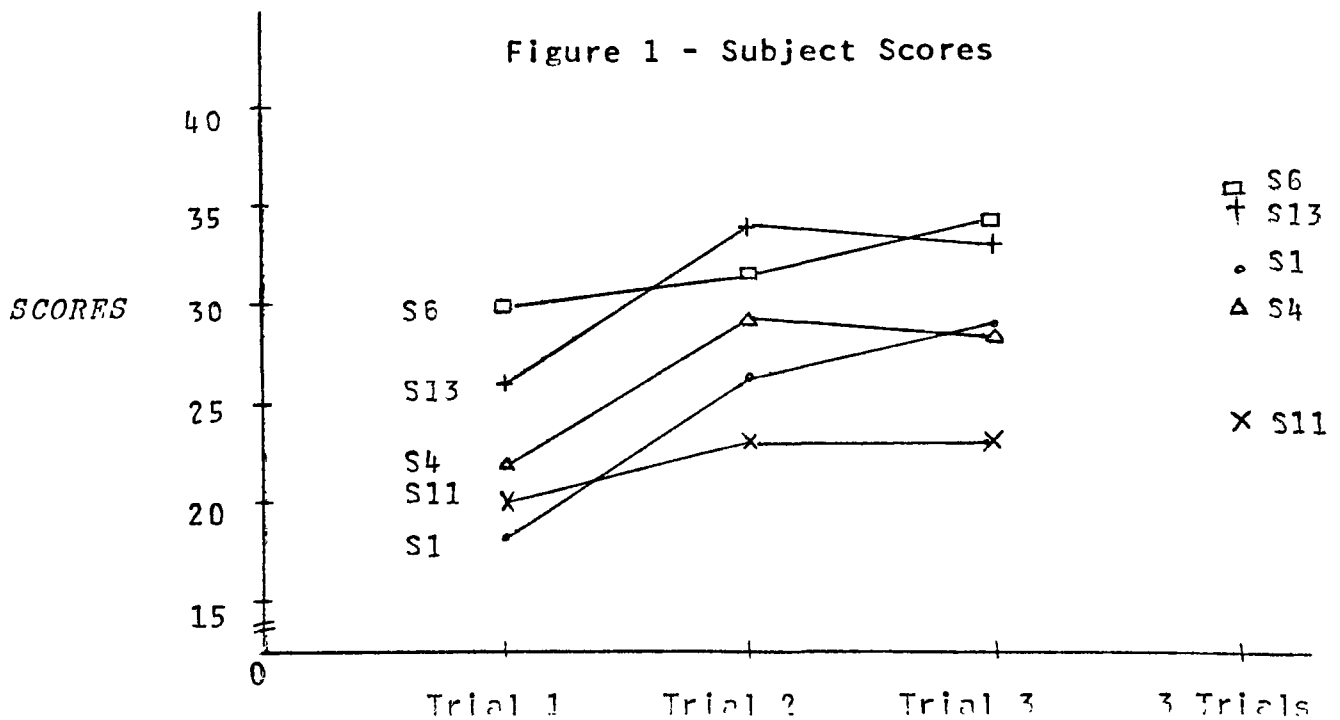
The four listening comprehension scores (Variables 1 to 4) are the number of questions completely correct (out of a possible 45) on various trials. Variables 1, 2, and 3 are the number of correct answers for each subject on the first, second, and third trials. Variable 4 is the number of answers which were correct on one or another trial. These are shown on Table 3. Also shown is the range of scores (difference between the highest and lowest score), the average score, the standard deviation, (these last two figures being results obtained from the computer output) and the standard error of the mean.

The most striking observation to be seen in this table is that great differences existed between subjects in listening comprehension as measured by this test. Scores range on the first trial from 18 to 30, or a range of 22 points, on the third trial from 24 to 36, or 12 points. S1 increased his score from 18 on the first hearing to 32 on the third, whereas S11 only improved from 20 to 24.

A second observation is that greater improvement is generally noted on the second trial than on the third trial.

Table 3 - Listening Comprehension Scores
Variables 1 - 4

	Trial 1	Trial 2	Trial 3	3 Trials
Subject 1	18	26	29	32
Subject 2	29	33	32	33
Subject 3	26	33	33	34
Subject 4	22	29	28	30
Subject 5	21	28	29	29
Subject 6	30	32	34	36
Subject 7	28	32	35	36
Subject 8	22	28	28	29
Subject 9	28	32	33	35
Subject 10	29	29	32	36
Subject 11	20	23	23	24
Subject 12	eliminated			
Subject 13	26	34	33	35
Subject 14	30	34	36	36
Subject 15	29	32	32	33
Range	22	11	13	12
Mean	25.5	30.4	31.2	32.7
St. Dev.	4.13	3.27	3.45	3.58
St. Error	1.14	0.91	0.96	0.99



c) Figure 1

Scores for subjects S6, S13, S4, S11, and S1, representative of some of the extreme and average scores, are plotted on a single graph for purposes of comparison and to illustrate better the first observation of the preceding section.

d) Figure 2

After all the questions were marked on the 5-point scale, (explained above in Ch. II Section 2.1), Figure 2 was constructed. It tabulates the data by subject (arranged vertically, S1...S15) and by question (arranged horizontally, Q1...Q45). Each column represents one question (nine questions on a page). Thus each square represents one subject's response to one question. It is divided into three columns which represent the three trials. The percentage of the answer correct is shown by the height of the shaded area in each column. The percentage represented by these heights is shown in the box at the bottom righthand corner of each page. This figure (Fig. 2) shows that for most of the incorrect questions, subjects still managed to understand with an accuracy of 50 to 75%. A at the bottom of the page represents the average score (for 14 subjects - S12 was eliminated) for each question; A1 the average on the 1st trial, A3 the average on the 3rd trial.

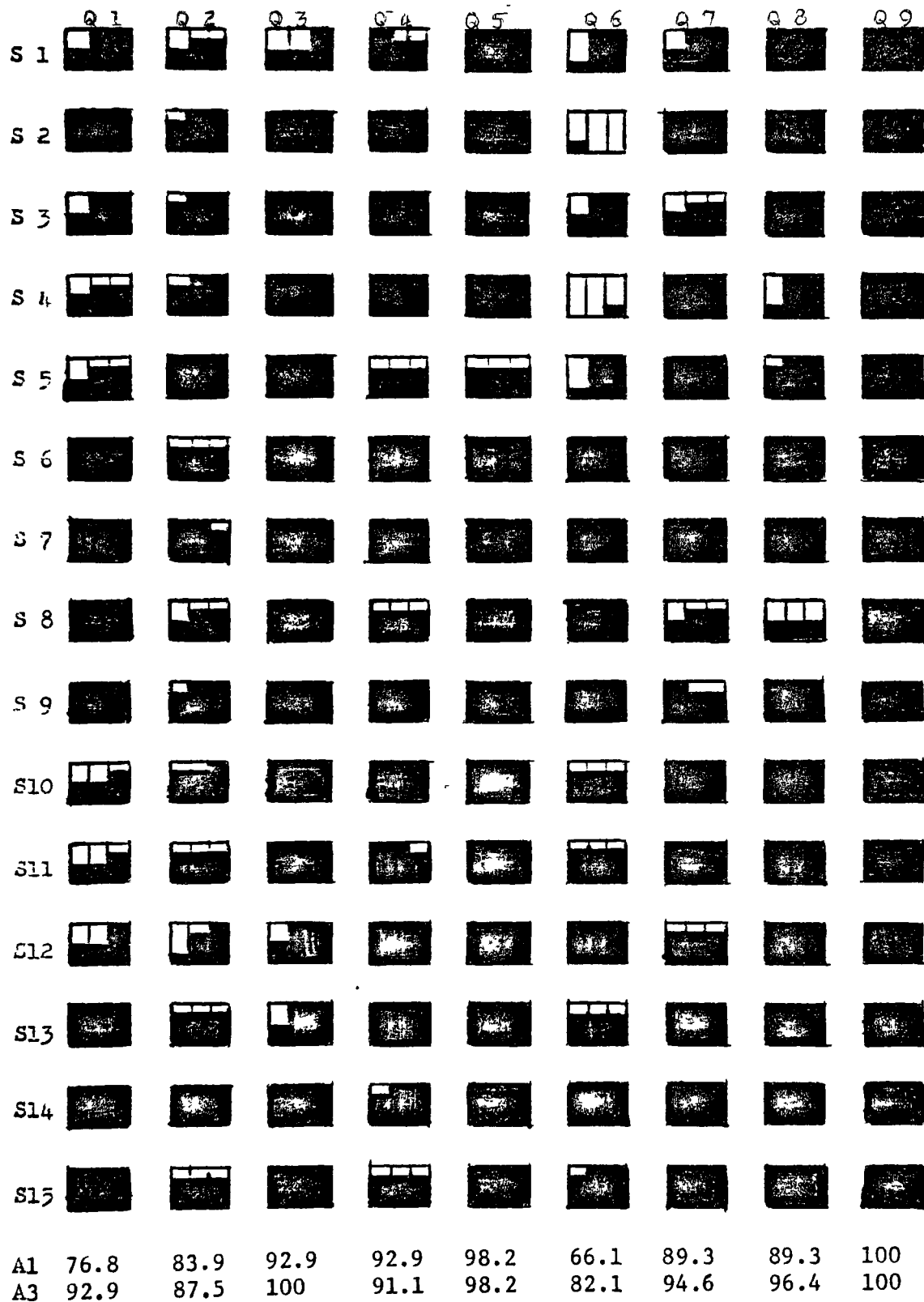
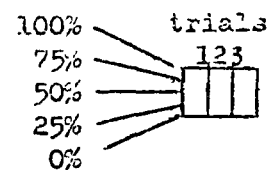


Figure 2



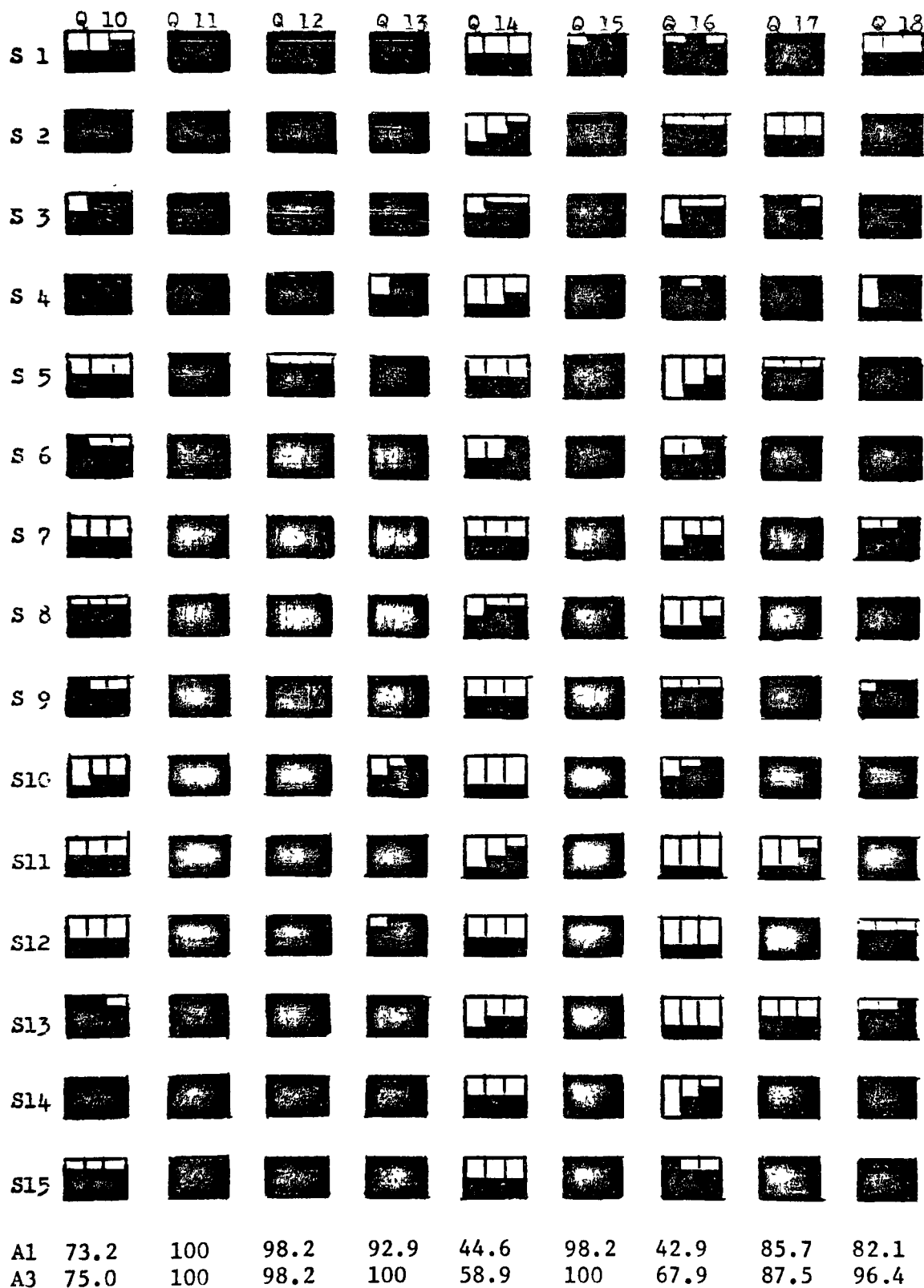
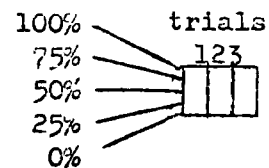


Figure 2 (cont.)



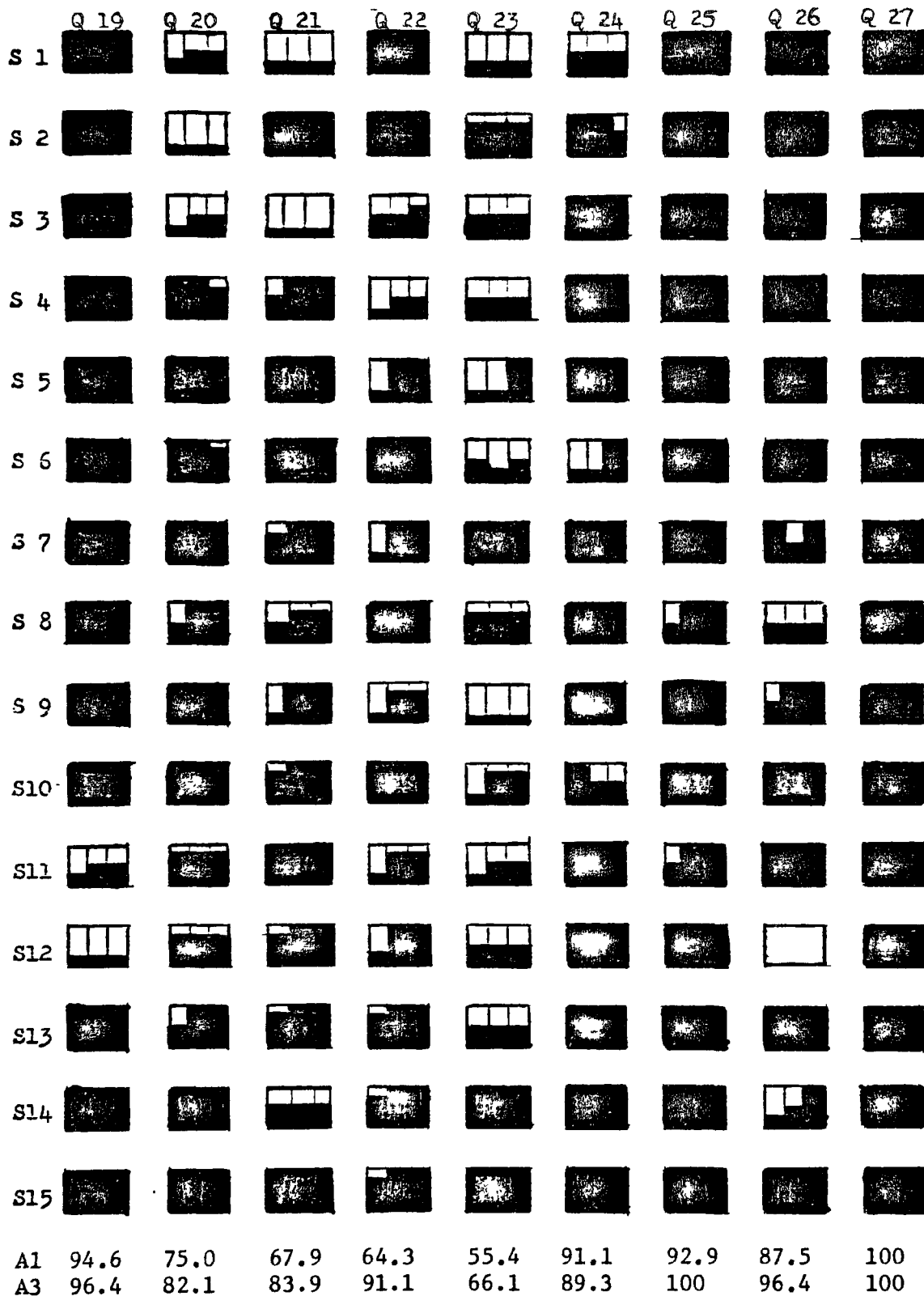
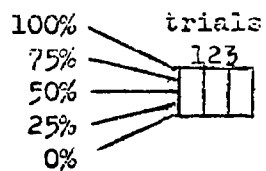


Figure 2 (cont.)



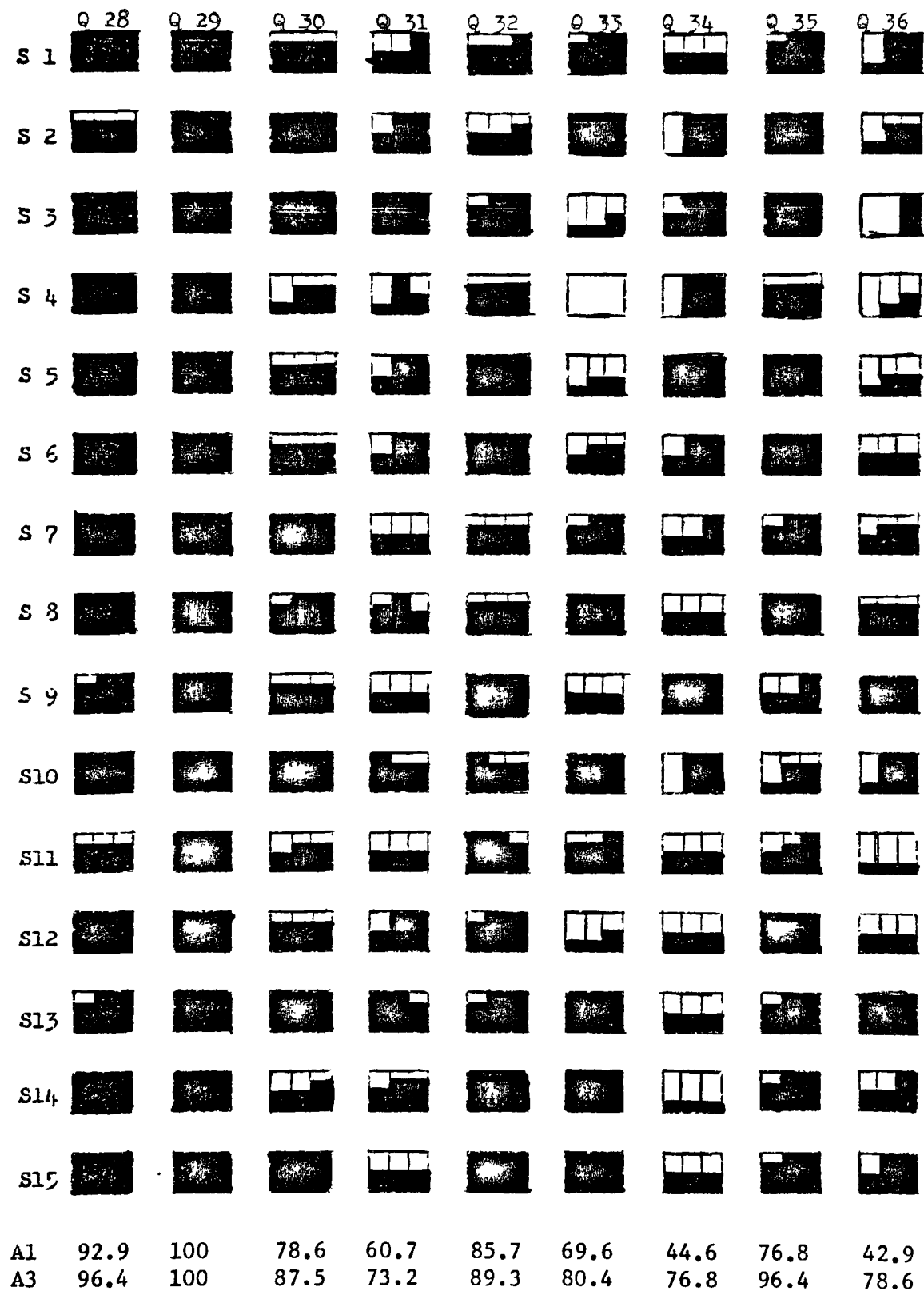
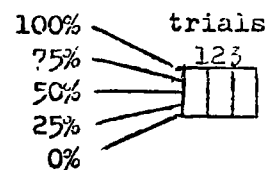


Figure 2 (cont.)



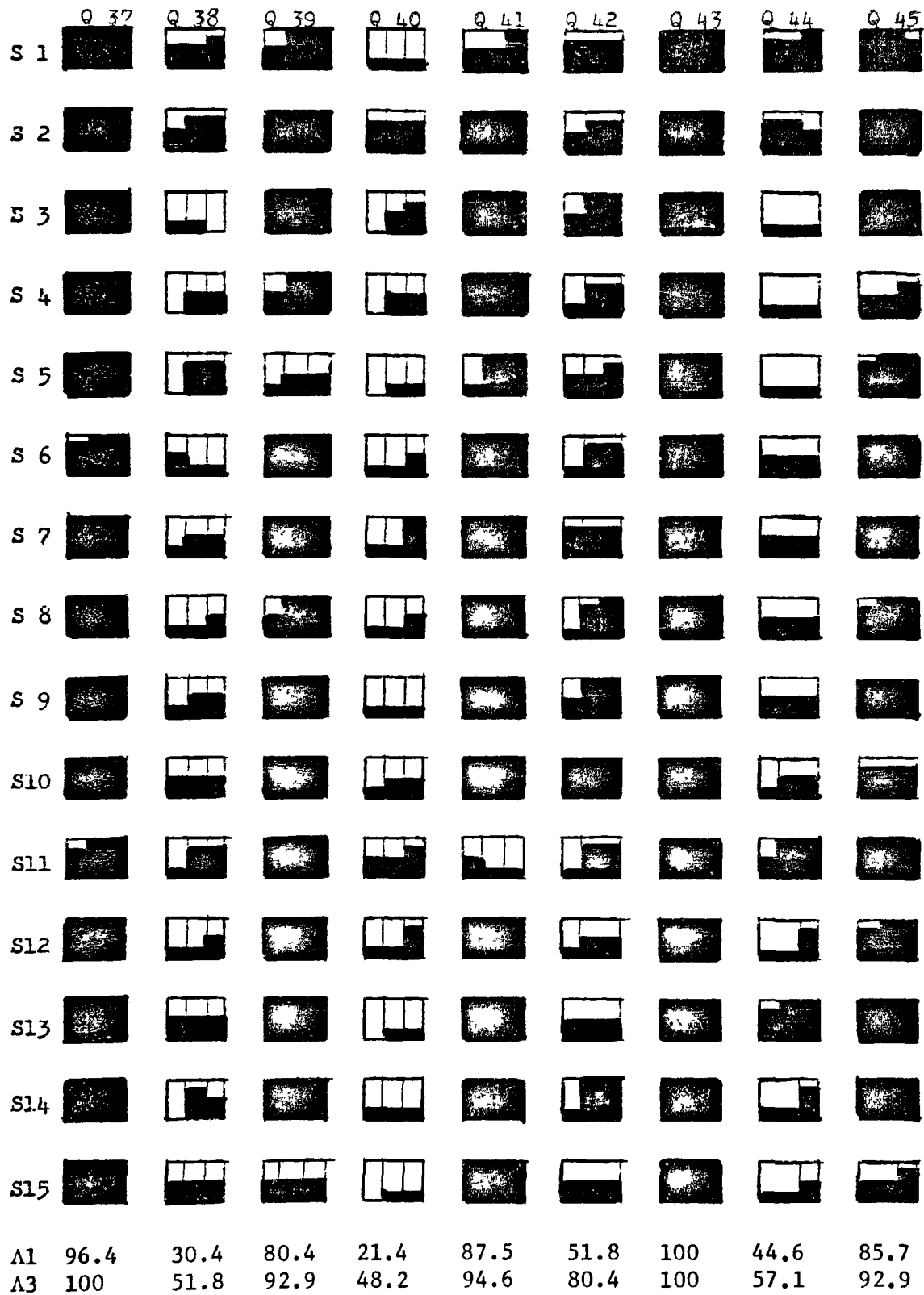
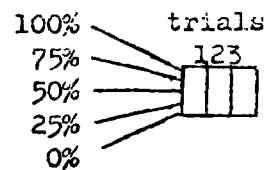


Figure 2 (cont.)



Tctal Score for Each Subject on First and Third Trials

Subject	Trial 1	Trial 3
S1	32	38 1/4
S2	37	40 1/2
S3	33 1/2	39 1/2
S4	29 1/4	37 1/4
S5	31 1/4	39
S6	37 1/2	40 3/4
S7	37 1/4	41
S8	34 1/2	39
S9	36 1/2	39 3/4
S10	35 3/4	40
S11	31 3/4	36 3/4
S12	eliminated	
S13	36 3/4	39 1/2
S14	36	41
S15	37 3/4	39 1/2

Figure 2 (cont.)

e) Table 4

The amount of improvement (Variables 5 and 6) made on each question was calculated. A gain is defined as a question on which any amount (25% - 100%) of improvement is observed on a given trial. The ratio of gains on a given trial to the number of mistakes made on the previous trial is the gain relative to mistakes. The sum of the relative gains on the second and third trials was Variable 5. Variable 6 is similar to Variable 5 except that it was calculated taking into account subsequent errors on the second and third trials. For example, a given subject may have been 75% correct on his first trial, but then have changed his answer to be only 25% correct on the second and third trials. Such "losses" were subtracted from the gains when V.6 was calculated.

$$V.5 = g(2)/m(1) + g(3)/m(2)$$

$$V.6 = e(2)/m(1) + e(3)/m(2)$$

where:

$g(N)$ = gains (number of questions receiving a higher score as explained above) on the Nth trial, as compared to the previous (N-1)th trial,

$e(N)$ = $g(N) - l(N)$, i.e. effective gain,

$m(N)$ = number of mistakes on the Nth trial,

$l(N)$ = number of losses (questions receiving lower score on Nth trial than before).

Table 4 shows this information in concise format. Here, for example, we see (reading across the first line)

that Subject 1 had 27 questions incorrect on the first trial, and made an improvement in 8 of these, giving a relative gain of $8/27$ or 40.7%. This left 19 questions incorrect on the second trial and he made an improvement on 7 of these, giving a relative gain of $7/19$ or 36.8%. Thus his total score for Variable 5, amount of improvement, was $40.7 + 36.8$ or 77.5 (shown in Totals column - V.5).

However, he had one loss between the first and second trials, reducing his first score to 35.2 for the first section, and two losses between trials 2 and 3, making his second score 26.3 for a total of $35.2 + 26.3 = 61.5$ for Variable 6, "amount of corrected improvement" (shown in Totals column - V.6).

Table 4 Listening Comprehension Improvement
(Variables 5 and 6)

Improvement:	Trials 1 - 2				
	m (1)	c (2)	g (2)	l (2)	e (2)
S1	27	11	40.7%	1	35.2%
S2	16	8	50.0	0	50
S3	19	12	63.2	0	63.2
S4	23	15	65.2	1	60.9
S5	24	14	58.3	0	58.3
S6	15	5	33.3	3	13.3
S7	17	8	47.0	1	41.2
S8	23	11	47.8	0	47.8
S9	17	8	47.1	2	35.3
S10	16	10	62.5	3	43.8
S11	25	11	44.0	1	40.0
S12	eliminated				
S13	19	10	52.6	0	52.6
S14	15	8	53.3	0	53.3
S15	16	5	31.8	1	25.0
Totals	272	136	50.0%	13	45.2%

Improvement:	Trials 2 - 3					Totals	
	m (2)	c (3)	g (2)	l (3)	e (3)	V. 5	V. 6
S1	19	7	36.8%	2	26.3%	77.5	61.5
S2	12	2	16.7	2	0	66.7	50.0
S3	12	4	33.3	2	16.7	96.5	79.9
S4	16	5	31.3	2	18.8	96.5	79.7
S5	17	3	17.7	0	17.7	76.0	76.0
S6	13	5	38.5	1	30.8	71.8	44.1
S7	13	4	30.8	1	23.1	77.8	64.3
S8	17	4	23.5	1	17.6	71.3	65.4
S9	13	1	7.7	0	7.7	54.8	43.0
S10	16	4	25.0	0	25.0	87.5	68.8
S11	22	6	27.3	2	18.2	71.3	58.2
S12	eliminated						
S13	11	1	9.1	2	0	61.7	52.6
S14	11	5	45.5	0	45.5	98.8	98.8
S15	13	2	15.4	0	15.4	46.7	40.4
Totals	205	53	25.9%	15	18.5%		
						Mean:	75.4 50.1
						St. Dev.:	15.6 16.7
						St. Error:	4.3 4.6

2. AUDITORY DISCRIMINATION TEST ANALYSIS (Variables 7 - 10)

Variable 7 derives from the first section of the Auditory Discrimination Test which presented pairs of tones where subjects had to discriminate pitch and intensity. Scoring of the ten questions was done on the basis of right or wrong. Variable 7 was the total score for these questions. Variable 8 derives from the second section of the Auditory Discrimination Test which required identification of synthetic vowels and is the total number of right answers in this section. A correlation easily became apparent in scoring this section, between scores on Variable 8 and Listening Comprehension scores. Subjects scoring higher on the Listening Comprehension Test tended to receive higher scores on recognition of synthetic vowels.

The third section, the presentation of short segments of the Listening Comprehension Test, was marked as to the number of the seventy phonemes correctly identified. Scores on the first trial were Variable 9, and scores on the second trial were Variable 10.

Table 5 shows all the subjects' scores for variables 7 through 10.

Table 5 Auditory Discrimination Scores
(Variables 7 - 10)

S	V.7	V.8	V.9	V.10
S1	8.5	5	35	45
S2	10.0	2	34	37
S3	10.0	7	35	34
S4	9.5	3	27	37
S5	8.0	1	36	42
S6	5.5	5	34	39
S7	9.5	8	38	47
S8	5.0	2	31	37
S9	9.0	5	37	38
S10	9.0	7	39	43
S11	8.0	4	27	34
S12	eliminated			
S13	7.5	7	43	46
S14	9.5	6	38	43
S15	9.5	5	35	42
Mean	8.46	4.8	34.9	40.3
St. Dev.	1.56	2.2	4.4	4.3
St. Error	0.432	0.596	1.21	1.19

3. HYPOTHESIS FORMULATION TEST ANALYSIS (Variables 11 - 16)

The first section on sentence endings was not considered in the tabulation and analysis of the data. It did not prove to be a good differentiating question; almost all the subjects attained perfect scores in this section.

The following section (Section 2) where subjects had to write down sixteen sentences read on tape provided three variables: (An answer was considered "acceptable" if it was a well-formed French sentence corresponding to the sounds.) The number of all acceptable answers when answers were marked right or wrong was Variable 11; the score when partially correct answers (marked on the 5-point scale explained in Listening Comprehension Test Analysis) were also added was Variable 12; the number of totally correct answers (i.e. answers which reproduced the speaker's intent) again marked right or wrong, was Variable 13. Table 6 shows the number of acceptable and totally correct answers for each question. It can be seen that in sentences 2, 10, 11, 15, and 16, acceptable answers and totally correct answers differ greatly. (See Ch. IV FINDINGS for a discussion of the errors subjects made in this section.) Some subjects were able to think of something other than the correct answer which would fit the sounds heard, or gave several acceptable answers. To measure this ability, a "degree of imagination" variable was calculated by taking a ratio of the number of unexpected but acceptable answers to the number of answers other than the one considered correct in

Variable 13. This was called Variable 15. Variable 16 was a sum of the scores of the second and third sections (Variable 11 and 14), as it was thought that perhaps together, they would act as a single hypothesis formulation variable. The questions in section three, requiring subjects to write down a second interpretation of the sentences printed on the answer sheet, were marked as to right or wrong. These scores are Variable 14. It was noted that two subjects showed a high degree of imagination (giving novel responses) in the third section as well as in the second section. In this former section, Subject 14 gave 4/10 unexpected answers; Subject 15 gave 3/10 unexpected answers. For example, for Q9 the printed sentence was "Laisse-les rentrer", the correct (and expected) answer was "Laisse l'air entrer". The novel answers (unique) given by S16 and S17 were "Les scellés rentrés" and "Lesley rentrait" respectively. However, the degree of imagination was not quantified in section 3 as other subjects did not give varying answers. It should be noted that S14 and S15 have high and moderately high scores on the Listening Comprehension Test respectively. Table 7 shows subjects' scores for Variables 11 through 16.

Table 6 Hypothesis Formulation Section 2 Responses

Question	Acceptable	Totally Correct
1. Est-ce que tout fnd vite?	8	8
2. L'homme à lui a son air. (ère, aire)	11	4
3. Tel fut l'hiver itère.	13	13
4. Pourquoi t'es-tu tu? (tes tutus)	14	14
5. Quel pou dur à tuer.	12	12
6. Certains-l'en.	7	7
7. eliminated		
8. Quand la laitue naît le ver sort.	14	14
9. Qui terre a, guerre a.	4	4
10. Etant (étang) mçu, il échoue. (est chou)	14	4
11. Jamais lard (l'art) n'alla mieux aux pcis. (poids)	13	3
12. Ce goût doux est pis que tout.	4	4
13. Mille ans de paix, c'est trop.	14	14
14. Qui dit kaki dit sale.	14	14
15. Ecoutez l'avis (la vie) air-terre.	7	3
16. Les en louer m'est dur. (les ans loués mais durs)	8	3

**Table 7 Hypothesis Permutation Scores
(Variables 11 - 16)**

S	V.11	V.12	V.13	V.14	V.15	V.16
S1	12	13.5	8	60	50	18
S2	12	12.8	10	80	40	20
S3	13	13.5	11	60	50	19
S4	12	13.0	8	50	57	17
S5	11	13.0	9	70	33	18
S6	8	10.2	6	80	33	16
S7	9	10.5	5	50	40	14
S8	11	12.0	7	40	50	15
S9	9	11.0	8	80	14	17
S10	11	12.0	7	70	50	18
S11	11	12.5	8	90	57	20
S12	eliminated					
S13	13	13.3	10	80	50	21
S14	12	12.8	9	90	80	21
S15	13	13.8	12	90	33	22
Mean	11.2	12.4	8.4	70.7	45.5	18.3
St. Dev.	1.58	1.14	1.9	16.4	15.4	2.4
St. Error	0.44	0.32	0.53	4.5	4.3	0.66

4. VARIABLES 17 - 19

The remaining variables used were:

Var. 17 : age in years.

Var. 18 : number of years away from France.
(calculated by subtracting the time spent on holiday in France from the number of years from the number of years away from France as indicated on the personal data sheets)

Var. 19 : sex of subject.
(masculine coded "0", feminine coded "1")

The amount of exposure to films was requested but was not counted as a variable because this information could not be obtained for all the subjects.

5. STATISTICAL ANALYSIS

Data was analysed by computer at the University of Ottawa Computer Centre and the Faculty of Education computer terminal and also by hand. Scattergrams and correlations for all the variables were made by the stock EMD 02D programme. (For correlations for all variables, see Appendix 3.)

Of the correlations, the following were significant at the .01 level of confidence: ($r \geq 0.6614$)

- V.3 : V.9 Listening comprehension score, third trial and Auditory Discrimination test segments, first trial
- V.4 : V.9 Listening comprehension score, three trials, and Auditory Discrimination test segments, first trial
- V.5 : V.15 Degree of improvement and degree of imagination
- V.6 : V.15 Degree of corrected improvement and degree of imagination
- V.14 : V.16 Hypothesis formulation section 3 and sections 2 & 3
- V.17 : V.18 Subject age and number of years in Canada

At the .05 level of confidence ($r \geq 0.5324$) the following were also significant:

- V.2 : V.8 Listening comprehension score, second trial and synthetic vowels score (Auditory Discrimination)
- V.3 : V.8 Listening comprehension score, third trial and synthetic vowels score (Auditory Discrimination)
- V.2 : V.9 Listening comprehension score, second trial and Auditory Discrimination test segments, first trial
- V.8 : V.9 Auditory Discrimination segments, second trial and synthetic vowels
- V.9 : V.19 Segments, first trial and sex (negative r)
- V.13 : V.19 Hypothesis formulation section two and sex (negative r)
- V.16 : V.19 Hypothesis formulation sections two and three together and sex (negative r)

Correlations were also computed between the hypothesis formulation variables 12, 13 and 14, and listening comprehension as measured by scores taken from Figure 2. For these correlations, partially correct answers were counted and added to the total score. A further correlation was calculated between listening comprehension variables 3 and 4, and the amount of exposure to films as measured by number of films watched per month. Neither correlation proved to be significant.

Since an interdependency was seen between Variables 3, 8, and 9, a multiple correlation was calculated by hand for Variable 3 correlated with Variables 8 and 9, which proved to be significant at the .01 level of confidence.

$$V.3 : V.8,9 \quad r = 0.76$$

Partial correlations were also calculated for Variables 3 and 9 removing the effects of 8, and for Variables 3 and 8 removing the effect of 9.

$$V.3 : V.9 (-V.8) \quad r = 0.608 \text{ (significant at the .05 level of confidence)}$$

$$V.3 : V.8 (-V.9) \quad r = 0.275 \text{ (not significant)}$$

The correlation between Variables 3 and 8 no longer remains significant. The correlation between Variables 3 and 9 is still significant at the .05 level of confidence.

The same procedure was followed for Variables 3, 17 and 18, since there was a high correlation between V.17 and 18, age and number of years away from France. A

multiple correlation between V.3, listening comprehension, Trial 3, and Variables 17 and 18 taken together was quite high, but not high enough to be significant for our sample.

$$V.3 : V.17,18 \quad r = -0.46$$

However, a partial correlation between V.3 and V.18 removing V.17 was significant at the .05 level of confidence.

$$V.3 : V.18 (-V.17) \quad r = -0.568$$

As the scattergrams (Appendix 4) indicate a number of curvilinear relationships, eta coefficients, or correlation ratios, were calculated by hand for some of the variables using the procedure given in Downie (1959, pp.116-119) and the scattergrams made by the computer. These figures are all high (in the range between 0.5 and 0.8), but they do not have the same levels of significance as do Pearson correlations. Tests for the significance of eta are to be found in Ferguson (1966, p. 249). These F-tests determine whether the eta obtained is significantly different from 0.0 or from linearity. The former test was computed using an APL program written by Dr. M. Cooper of the Faculty of Education based on the Ferguson formula.

$$F = \frac{\eta^2 / (k - 1)}{(1 - \eta^2) / (N - k)}$$

where:

N = number of cases

η = eta coefficient

K = number of arrays, either rows or columns

Using this formula and the F-table (Ferguson, p. 408) none of the eta scores were found to be significantly non-zero. If the number of subjects had been greater, it would have reduced the magnitude of the F value needed to reject the null hypothesis and perhaps produced significant results.

Spearman's coefficient of rank correlations or Spearman's r_s were also calculated for these variables, using one of the Faculty of Education APL programmes. Significant correlations obtained were: (1) .05 level of confidence [$r \geq 0.456$], (2) .01 level of confidence [$r \geq 0.645$])

- V.2 : V.8 $r = 0.485^{(1)}$ Listening comprehension score, second trial and synthetic vowels
- V.3 : V.8 $r = .692^{(2)}$ Listening comprehension score, third trial and synthetic vowels
- V.4 : V.8 $r = 0.644^{(2)}$ Listening comprehension score, three trials and synthetic vowels
- V.3 : V.9 $r = 0.657^{(2)}$ Listening comprehension score, third trial and segments
- V.4 : V.17 $r = 0.573^{(2)}$ Listening comprehension score, three trials and age

A test of the significance of the difference between the means (t-test) was made for Variables 11, 12, 13, the three different scores of the hypothesis formulation test. The results (shown in the following Table 8) show that these means are significantly different.

Table 8 Significance of the Difference Between Means

V	Mean	St. Dev.	N	t
11	11.21	1.58	14	2.32 ⁽¹⁾
12	12.42	1.14		
12	12.42	1.14	14	6.72 ⁽²⁾
13	8.43	1.91		
11	11.21	1.58	14	4.23 ⁽²⁾
13	8.43	1.91		

(1) significant at the .05 level of confidence

(2) significant at the .01 level of confidence

CHAPTER IV

FINDINGS

1. LISTENING COMPREHENSION TEST

1.1 First of all, the results show that native speakers can usually understand isolated utterances without the aid of situational or linguistic context, except as provided by the sentence itself. Understanding is often attained in spite of an extremely poor acoustic signal and to a high degree of accuracy. Only three questions proved difficult for all subjects. This was contrary to our expectation that most, if not all, of the forty-five utterances selected for the test would present difficulty for the subjects. Explanations may be found in the following facts: A native can use his knowledge of what is possible in a given grammatical (e.g. a verb terminated by 'pas' usually follows the combination "je ne le " nl) or semantic environment, as a given utterance provides some linguistic context for each of its component words. The native may also use the semantic environment to limit the possibilities for an unknown word. Also playing a role may be cultural factors, such as knowing what stock expression¹ is likely to be used in any given situation. One would expect (see Savir, 1962) that frequency of usage may also be important for the recognition of given words. It is virtually impossible for most non-natives to understand many of the utterances of the listening comprehension test which were understood by natives. The better scores of native speakers may, however, be partly attributable to a higher motivation "to perform as native speakers ought to".

Responses, even when mistaken, were not only similar phonetically and rhythmically to the correct answer, but also tended to be grammatical sentences. Non-acceptable ungrammatical sentences received as answers from the 14 subjects accounted for only 21 of the total 1890 answers, or a mere 1.1%. For example, for Q.16, which was "les médecins ragent, mon vieux", subjects gave responses such as "il est de son âge", "l'enreni de son âge", "les maux de son âge". Q.6 "Ça te plaît moins que la perche" produced answers such as "appelez-moi Claperche". This substantiates Thorne (1966, p.8) who notes that badly-formed sentences are heard as well-formed sentences, and that sentences other than those given are heard in noisy conditions. He says that some features of speech are not present in the acoustic signal itself, but are supplied by the listener. (p. 5)

1.2 Second, our finding that native speakers show great differences in ability (observed as early as 1926 by Hankin) to hear and comprehend utterances such as those used in our listening comprehension test casts doubt on the use of the term "native, or near-native ability" often cited as one of the goals of foreign language methods. Writers of text books for language learners have often defined their goal as native speaker competence. For our test data, there was no observable clustering of native-speaker competence; ability in these circumstances varies considerably. On this point, Perren (1968, p. 111) comments,

"But in language teaching and language testing the criterion is usually assumed to be that of native speakers without further definition. Terms such as 'native-like skill' are commonly used at the top end of rating scales, but what exactly do they mean?"

Studies (e.g. Bird, 1953) show that in listening comprehension natives also can improve with training. It seems clear that the goal of second (or foreign) language learning should be defined, but even supposing that native-like ability could be adequately defined (which our research throws into doubt) it is probably not a very realistic goal in a country or area where the language is not commonly spoken. If anything, it is rather utopian. It is questioned in particular whether the goal of trying to teach students a high level of listening comprehension skill, the ability to understand films, plays, and other material where there is no face-to-face contact, no chance for a question or repetition, is at all feasible for students in particular areas.

It is not clear how such a goal can be achieved. Our results show that most improvement in listening comprehension occurs on the second hearing. A third hearing provides little improvement or even loss of comprehension. Table 4 "Listening Comprehension Improvement" shows that on Trial 2, average improvement was 50%, i.e. 50% of all items heard incorrectly on Trial 1 were improved, whereas on Trial 3, improvement was half as great: only 26% of items heard incorrectly on Trial 2 were improved. Although the total number of losses remains about the same on both Trials 2 and

3, the number of gains falls sharply so that a change on trial 3 has a greater chance of being a loss. (The ratio of losses to gains on Trial 2 = $13/136 \cong 10\%$, on Trial 3 = $15/53 \cong 29\%$.) Half the subjects showed a greater number of losses on the third trial than on the second trial, and about one-third showed no change.

In the research undertaken previous to this experiment, the experimenter listened to various selections of about 100 words before playing isolated utterances of these selections to subjects. It was found that after three tries at understanding an utterance in a foreign language further attempts did not lead to improved understanding, as the same divisions of sound were made repeatedly. To illustrate this, we give an example from the Listening Comprehension Test data, where one of the subjects made his first division incorrectly, and then was unable to hear anything else. (18, "Cù dessine encore Rembrandt" was interpreted as "Ncus dessincns - cor - Rembrandt" with the isolated nonsense syllable "ccr" in the middle. It remained unchanged during subsequent trials.

1.3 Blanks and Hypotheses

During the analysis of the data, it was perceived that subjects who left blanks on the preliminary trials (1 and 2) tended to fill them in correctly on later (2 and 3) trials, while subjects who made a wrong hypothesis tended to keep it. Since this tendency has not been verified for

listening comprehension, and having the data to examine it, we decided to extract the pertinent statistics about this observation.

We looked only at the differences between the first and second trials, since this was where most improvement was made. Also scoring becomes complicated with three trials when a subject has left a blank on the first trial, made a hypothesis on the second and given a correct answer on the third, or other combinations of blanks and hypotheses. Since this hypothesis was formulated and tested independently of the rest of the analysis, we counted each question as many times as there were different occurrences of hypotheses and blanks which satisfied the criteria. To illustrate - a blank which was partly filled in correctly and partly left blank was counted as two blanks, one which became right (Br) and one which remained blank (Bb). Also cases where the subject was unaware of a portion of the utterances (i.e. blanks were not left by the subject, even though words [or a word] were left out) were not counted as blanks, as not being aware of something missing would not usually lead to an improvement in that portion of the utterance.

The hypotheses were formulated in the following way:

Hw = the number of occurrences of a wrong hypothesis on Trial 1 on the Listening Comprehension Test which was changed on Trial 2, but still not completely correct.

H_r = the number of occurrences of a wrong hypothesis on Trial 1 which was changed to a correct answer on Trial 2.

H_h = the number of occurrences of a wrong hypotheses which remained unchanged on Trial 2.

H = the number of wrong hypotheses on Trial 1. ($H = H_w + H_r + H_h$)

- a) The number of wrong hypotheses which remained unchanged would be greater than the other types of hypotheses ($H_h > H_r$ & $H_h > H_w$), so that the percentage of unchanged hypotheses over the total number of hypotheses would approach 100% ($H_h/H \rightarrow 100\%$).
- b) Few hypotheses would result in correct answers. That is, the percentage of hypotheses which have been revised correctly over the total number of hypotheses would approach 0% ($H_r/H \rightarrow 0$).

E_w = the number of occurrences of a blank on Trial 1 which was filled in incorrectly on Trial 2.

E_r = the number of occurrences of a blank on Trial 1 which was filled in correctly on Trial 2.

E_b = the number of occurrences of a blank on Trial 1 which remained a blank on Trial 2.

E = the total number of occurrences of a blank on Trial 1. ($E = E_w + E_r + E_b$)

- c) Most blanks on Trial 1 would be filled in correctly so that the percentage of these blanks over the total number of blanks would approach 100% ($E_r/E \rightarrow 100\%$).
- d) Relatively speaking, more hypotheses would remain unchanged than blanks would, that is, the percentage of H_h would be greater than the percentage of E_b ($H_h/H > E_b/E$).

The finding was that for hypotheses, a) could be supported whereas b) could not. The ratio of H_h over the H was $121/200$, i.e. 61% of wrong hypotheses on Trial 1 remained unchanged on Trial 2.

The ratio of H_r over H was $63/200$, i.e. 31% of all wrong hypotheses on Trial 1 were revised better on Trial

2. However, Hr did not turn out to be the smallest of the three outcomes, but was half as great as Hh. The ratio of Hw over H has 16/200 or 8%, surprisingly the lowest of the three outcomes.

For the blanks, results were not as expected. The number of Br was about the same (30/107) as the number of Bb (32/107) and the number of Bw (45/107) was about 50% more than the first two scores. Of the blank answers on Trial 1 which did not remain blank on Trial 2 (107 - 32 = 75), 40% were replaced by completely correct answers, but all of the remaining 60% cannot be said to be completely wrong; most were improved from the first trial, but since our court only considered completely correct answers for Br, even one incorrect phoneme made an answer Bw. Therefore c) was not supported numerically by this analysis. However, d) comparing the percentages of hypotheses and blanks unchanged is supported by the data which says twice as many hypotheses remained unchanged as blanks (60% to 30%). See Table 9.

Table 9 Blanks and Hypotheses

S	Hw	Hr	Hh	H	Bw	Br	Bb	B
S1	2	7	15	24	2	2	0	4
S2	0	5	8	13	2	2	0	4
S3	3	6	4	13	3	5	5	13
S4	0	6	6	12	6	6	4	16
S5	0	3	9	12	6	4	4	14
S6	3	4	8	15	1	0	1	2
S7	0	2	8	10	4	2	1	7
S8	2	6	10	18	5	1	1	7
S9	0	5	9	14	2	3	3	8
S10	1	2	9	12	2	2	1	5
S11	2	5	10	17	7	2	5	14
S12	eliminated							
S13	0	7	9	16	2	0	5	7
S14	2	1	7	10	2	1	1	4
S15	1	4	9	14	1	0	1	2
Totals	<u>16</u>	<u>63</u>	<u>121</u>	<u>200</u>	<u>45</u>	<u>30</u>	<u>32</u>	<u>107</u>
Percent- ages	8%	31%	61%		43%	28%	30%	
Total					75			
Percentages					60%	40%		

1.4 Error Analysis - Listening Comprehension Test

An analysis of errors ("semantic" errors, where the syntax has been correctly identified) shows that when items are erroneous, it is because of noise² in the channel combined with factors of frequency and semantic acceptability; that is, when faced with a choice of several words, subjects tend to choose a word which is most frequent and also likely to occur in a given context. For example, half the subjects did not perceive the clear /g/ of the name "Greux" (Q.2), a small unknown French village, but heard "Dreux", which is a well-known town. Cultural expectancy may also be playing a role here. Most changed to Greux after the second or third trial, but a couple maintained Dreux throughout the three trials. Other mistakes for this rare, "Grue, Gru, Greuse", indicate that although syntactically all the subjects realized a place name had to occur in that location, they were unable to hear the vowel correctly. In Q.6 the replacement of "perche" by "pêche" can also be attributed to the factor of frequency.

It has been shown that in a situation without a clear stimulus, natives will hear something that is phonemically quite close and more frequent (Savin, 1962). This word-frequency knowledge seems to be something that is unconscious and acquired by every native and may be difficult to teach as a skill in foreign-language learning.

In Q.16 the expression "mon vieux" was

interpreted as "mon Dieux" by about half the subjects while a couple of subjects switched from one to the other and back again in the course of the three trials. "Mon Dieux" seems to be more accurate, as repeated listenings seem to indicate more of a /d/ than a /v/³, but from the context of the film and from the prosodic features, it cannot be other than "vieux". It seems that the prosodic aspects of utterances, cited as being one of the keys to listening comprehension (Lieberman, 1967) are often over-ridden by natives. The previous example for Q.6 "Ca te plaît moins que la perche?" is a question, whereas the answer "Appelez-moi Claperche" is a command normally with a different intonation. Furthermore, in the second subsection of the Hypothesis Formulation Test, the last question "Les en louer m'est dur" was interpreted as "les ans loués mais durs" by one third of the subjects in spite of the intonation of the sentence. Because the French native who read this and all the other sentences in this section did not know this other possibility, it is possible, but highly unlikely, that he did pronounce something ambiguous.

Semantic incompatibility is another source of errors.

Q.31 Socrate était donc un chat.

Subjects had difficulty hearing the assertion of Socrates being a cat and came up with 'saint' /sE/ or 'chief' /Ef/ for 'cat' /a/, or 'this crab' /s krab/ and 'this ace' /s krak/ for 'Socrates'.

Ç.10 Eh ben, pas si bête mon cher.

This was made by two-thirds of the subjects into a more complete 'Elle est ben pas si bête' or changed to 'Va pas si bête, mon cher' or 'Va pas si vite, mon cher'. 'bête' was changed to 'vite' because it is more compatible with 'va'. Subjects apparently override the acoustics of the message in their attempt to make some kind of sense out of what they hear.

2. HYPOTHESIS FORMULATION TEST

2.1 Error Analysis

Questions 9, 12, and 6 proved to be the most difficult questions as they had the fewest right answers, Q.9 and Q.12 with four right answers, Q.6 with seven.

Q.9. Qui terre a, guerre a.

Q.9 presented an unusual syntactic pattern, of NP + NP + V, NP + V which is not found outside of set expressions, such as proverbs.

Q.12 Ce goût dcux est pis que tout.

/Epik/ could not be segmented into three words, "est pis que", by many. It was interpreted by some as one verb "épique", by others as the verb "est" plus adjective "pique". The word "pis" is usually found only in set expressions such as "tant pis", "il a dit pis que pendre de vcus", but is not known to native speakers. All the subjects who were told the correct answer afterward immediately recognized the correctness of this segmentation without further explanation.

Q.6 Convaincs-l'en.

In Q.6 the combination of the two pronouns "le + en" proved difficult to hear. It does not seem to be a common combination. Some heard only one pronoun "la", others divided "convaincs" into "qu'en vint".

Questions which were correct but with unexpected answers were Numbers 2, 10, 11, 15, and 16.

Q.2 L'homme a rui à son air.

The majority of answers contained the word "ère" for /Er/, a much less frequent word than "air" and it seems to make a much less likely sentence: 'Man has harmed his era' instead of 'air'. One explanation may be that "air" is not usually found with a possessive; "l'air" is the more common form, found in such expressions as "avoir, prendre l'air, en l'air". Therefore, because "son" was used, subjects tended to think of 'era', found in the expression "de notre ère". One subject wrote "aire".

Q.10 Etant mou il échoue.

The results of this question were surprising. The meaning of the sentence was that, since "he" was weak-willed, he was failing; instead, about half the subjects wrote "il est chou" or 'he is a darling'. One subject interpreted /eta/ as "étarg" or 'pond'.

Q.11 Jamais lard n'alla mieux aux pois.

This was a French proverb where subjects gave alternate answers for /lar/ and /pwa/ meant to be "lard" and "pois" or 'bacon and peas'. Here answers were, "l'art", also Larre (a proper name?) and "poids" or 'weight'.

Q.15 Ecoutez l'avis air-terre.

Here /lavi/ was interpreted as "la vie" or 'life' instead of "l'avis" or 'advice'. A combination of three nouns, such as "l'avis air-terre" is understood with difficulty if it is not a frequently-used combination such as "le match France-Allemagne" or "les réunions patrons-ouvriers". For "air-

terre" subjects wrote proper names such as "Herther,
Erther".

Q.16 Les en louer m'est dur.

Subjects interpreted /lwe/ as "loué" or "louez" which were not acceptable answers and /mE/ as "mais", in some cases producing the acceptable answer, "Les ans loués mais durs", which, however, as noted before, should have a different intonation pattern.

3. CORRELATIONS

The most significant correlations obtained were between listening comprehension and the segments of sound section of the Auditory Discrimination Test. Although the synthetic vowels section seemed to correlate with listening comprehension was shown to correlate as a result of its relation to the segments section. When a partial r was calculated with the effects of the other section removed, synthetic vowels no longer correlated. Listening comprehension did not correlate with the age of the subject, meaning younger subjects did not do significantly better than older ones. However, partial correlations show that listening comprehension seems to be related to number of years subjects have been away from France; all the subjects had been away from one to four years. We tested another subject who had been away thirteen years for comparison with our subjects. This subject's score on the Listening Comprehension Test was much lower than for the other subjects, indicating that listening comprehension may decline when this ability has not been used for a long time.

There was no significant correlation of listening comprehension with hypothesis formulation. Interestingly, there was a correlation between the degree of improvement made on subsequent trials of the Listening Comprehension Test and the degree of imagination calculated for the Hypothesis Formulation Test, section 2. The sex of the subject correlated negatively with auditory discrimination,

Section 3, and with the hypothesis formulation scores of sections 2 and 3 combined, meaning that males did better than females in these sections.

Significant-seeming scattergrams were obtained for some of the variables as shown in Appendix 4. However, for the number of subjects we had, none of them turned out to be statistically significant.

Notes to Chapter IV

¹ Stock expressions form a great part of our everyday speech, so much so that deviations from them are understood only with difficulty. See the discussion of these in connection with the hypothesis formulation section 2 in the following Error Analysis - Hypothesis Formulation Test.

² It is never possible to remove all noise from a listening situation

³ Spectrographic analysis of the utterance did not prove helpful because of the interference of noise.

CHAPTER V

CONCLUSIONS

1. GENERAL CONCLUSIONS

The following conclusions were derived for subjects listening to rapid conversation in their native language. (These are discussed in greater detail in the previous chapter.) Auditory discrimination, especially as measured by the subtest consisting of segments of sound, correlated with listening comprehension. Hypothesis formulation ability, as defined by our testing procedure, did not. Listening comprehension also correlated negatively with number of years away from France; recently-arrived subjects tended to score higher on the Listening Comprehension Test. Listening comprehension did not correlate with the amount of time subjects spent watching films. Also significant was the correlation between the degree of improvement made by subjects from one trial to another on the Listening Comprehension Test and the "degree of imagination" factor in the Hypothesis Formulation subtest. See Chapter III Section 1.e) Listening Comprehension Test Analysis and Section 3. Hypothesis Formulation Test Analysis for a detailed description of these measures.

There was also a noticeable tendency for subjects scoring high on the Listening Comprehension Test to display greater creativity in their hypothesis formulation answers. See Section 3, Chapter III.

CONCLUSIONS STATE IN NULL HYPOTHESES:

(In the following list of null hypotheses, (+) marks a hypothesis kept and (-) marks a hypothesis rejected at the .05 level of confidence.)

1. Listening comprehension is not related to auditory discrimination.

(+) Part 1 tones

(+) Part 2 synthetic vowels

(-) Part 3 segments

(+) 2. Listening comprehension is not related to hypothesis formulation.

(-) 3. Listening comprehension is not related to age.

(-) 4. Listening comprehension is not related to sex.

(+) 5. Listening comprehension is not related to number of years away from France.

(+) 6. Listening comprehension is not related to amount of exposure to films.

(-) 7. There is no relationship between auditory discrimination and hypothesis formulation.

2. IMPLICATIONS

It was observed that a higher hypothesis formulation score was made by subjects who scored lower in auditory discrimination, so there appears to be a slight tendency toward compensation working between these two factors. That is, persons lacking in auditory discrimination ability have greater ability in hypothesis formulation and vice versa.

In second-language teaching, one tends to assume a certain level of listening comprehension ability when hearing a student talk. This experiment has shown how widely listening comprehension abilities can range in native speakers. Therefore, oral production is far from definitive in the evaluation of listening comprehension for fluent, native speakers.

There was an unexpected disparity of scores on the Listening Comprehension Test. For successful teaching, the goals of teaching must be considered and established objectively. This apparently cannot be done in terms such as "native-like" or "near-native" skill. This experiment shows that there is no clustering of native ability in listening comprehension. Anyone trying to define such terms is apparently engaged in a hopeless task.

The scattergrams show that males tend to do better than females both in the Listening Comprehension and Hypothesis Formulation Tests. This supports Nichols' (1948)

and Caffrey's (1955) findings.

3. SUGGESTIONS FOR FURTHER RESEARCH

This experiment raises some questions for the teaching and testing of listening comprehension in a second language. First of all, to what level of proficiency should teachers and methods aim? Most students are able to reach the level of understanding a simple face-to-face conversation. The difficulty lies in the area between that stage and the stage of being able to follow films and television programmes easily. Perhaps there should be two programmes of studies or an optional alternative that the student may take, depending on whether he wants to achieve a second, higher level of comprehension or not. But it can even be questioned whether ability to understand utterances such as those used in our Listening Comprehension Test is a feasible goal for even advanced students.

What drills are available for teaching an advanced level of listening comprehension? Most exercises available at present consist of passages to be read orally to a class with questions afterwards to test comprehension. Other drills in addition to the minimal-pair drill must be developed for auditory discrimination. A possible exercise might be to have a student listen to a sentence while he reads a printed transcription, and to ask him to modify the transcription to correspond to what he has heard.

The poor correlation of hypothesis formulation ability with listening comprehension may have been due to the design of the test. Further study is needed on techniques for testing this. Another possible study might be to see if listening comprehension improves after training in hypothesis formulation exercises such as the following:

- a) Giving a student the ending or beginning of a word and requiring him to find all the possible words with that ending or beginning.
- b) Giving a situation and topic and asking the student to think of all the words, for example, with two syllables or ending in "ere" related to the given situation.
- c) Playing a passage with words cut out and asking the student to supply the words which would make sense in the spaces.

Most of the questions raised could not be answered definitively because of the small number of subjects. For reliability with a small sample, very high correlations are required, and factors which are only poorly correlated in reality cannot be claimed as significant. For example, it is accepted that auditory discrimination decreases with age. However, our correlations showed only a slight negative relationship between auditory discrimination and age and the size of the sample needed to show this to be significant would have been between thirty and thirty-five subjects.

Also, the tendency for males to do better is not in the scope of the thesis to investigate further. This finding should perhaps be verified with a larger population, and if it still holds true, a study ought then be made to see what traits in males make for better listening comprehension.

With hindsight, we make the following suggestions if the experiment were to be replicated. We would control for IQ more strictly with some kind of intelligence test, although it has been questioned whether intelligence is a factor in listening comprehension (Escott, 1973, p. 30). (Our experiment used only a weak control for intelligence; subjects had roughly the same educational level.)

A better measure of expectation of the usual is needed, as this seems to be an important factor in listening comprehension. Subjects had difficulty comprehending unusual, novel combinations of words, even when they were clearly pronounced. Most of the utterances we hear in our native language are habitual, cliched, and expected in certain situations.

For auditory discrimination, perhaps the difficulty in the section of sound segments was the rapid changes of speakers. Perhaps modifications could be made to accustom the listener to each speaker's voice before each segment.

Bibliography

- Abercrombie, D. et al. (Ed) In Honour of Daniel Jones (1964) London (Lergmans, Green & Co.).
- Ainsworth, S. & Figh, C. "Auditory Functions and Abilities in Good and Poor Listeners" (1954) JComm 4: 84-6.
- Fird, D. "Teaching Listening Comprehension" (1953) JComm 3 (2): 127-30.
- Flack, C. "Graded Practice in Advanced Listening Comprehension" (1971) English Language Teaching 25 (2): 171-8.
- Flack, J. "Predicting the Intelligibility of Words" (1961) Proc. 4th Int. Cong. on Phon. Sci., The Hague (Mouton) pp. 347-54.
- Elewett, Th. "An Experiment in the Measurement of Listening at the College Level" (1951) JComm 1 (1): 50-7.
- Erière, E. "An Experimental Approach to Defining a Hierarchy of Difficulties of Learning Phonological Categories" (1966) Language 42: 768-96.
- Ercadent, D. "Attentive and the Perception of Speech" (1962) ScAm 206 (4, April): 143-51.
- Ercwn, J. "The Objective Measurement of Listening Ability" (1951) JComm 1 (1): 44-6.
- Ercwn, J. "The Measurement of Listening Ability" (1950) School and Society 71: 69-71.
- Erown, C. & Rubenstein, h. "Test of Response Bias Explanation of Word-Frequency Effect" (1961) Science 133: 280-1.
- Eruce, D. "Effects of Context upon Intelligibility of Heard Speech" (1956) in Cherry, C. (Ed) Information Theory: 3rd London Symposium (Butterworth) pp. 245-252. (1968) Reprinted in Oldfield & Marshall pp. 123-131.
- Eruce, D. "The Effect of Listener's Anticipation on the Intelligibility of Heard Speech" (1958) Lg&Sp 1: 79-97.
- Burger, S. "Listening Comprehension: Its Components and a Model for Testing" (1972) M.A. Thesis (Univ. of Ottawa).
- Caffrey, J. "Auding" (1955) REdRes 25: 121-38.
- Cartier, F. "Listenability and 'Human Interest'" (1955) SpM 22: 53-7.
- Cross, D. et al. "Identification and Discrimination Functions for a Visual Continuum and their Relation to the Motor Theory of Speech Perception" (1965) JExpPsy 70 (1): 63-74.
- Delattre, P. "Le mot est-il une entité phonétique en français?" (1966) in Delattre, P. pp. 141-9.
- Delattre, P. Studies in French and Comparative Phonetics (1966) The Hague (Mouton & Co.).
- Delattre, P. "Syntax and Intonation: A Study of Disagreement" (1970) MIJ 54 (1): 3-9.
- Denes, P. & Pincsn, E. The Speech Chain; the Physics and Biology of Spoken Language (1963) Baltimore (Bell Telephone Laboratories, Inc., The Williams & Wilkens Co./ Science Series) pp. 125-47.

- Downie, N. & Heath, R. Basic Statistical Methods (3rd, 1970) New York (Harper & Row).
- Erickson, A. "Can Listening Efficiency be Improved?" (1954) JComm 4 (4): 128-33.
- Escott, M. "Are You Listening?" (1973) En Route 1 (4, Apr): 27-30.
- Fairbanks, G. et al. "Effects of Time Compression upon the Comprehension of Connected Speech" (1957) JSHD 22: 10-19.
- Ferguson, G. Statistical Analysis in Psychology and Education (2nd, 1966) New York (McGraw-Hill).
- Finocchiaro, M. English as a Second Language (1964) New York (Regents).
- Fodor, J. & Bever, T. "The Psychological Reality of Linguistic Segments" (1965) JVLVB 4: 414-20.
- Fry, D. "Experiments in the Perception of Stress" (1958) Lg&Sp 1: 126-52.
- Fry, D. "Speech Reception and Perception" (1970) in Lyons, J. (Ed) New Horizons in Linguistics, Hammondsworth & Baltimore (Penguin Books) pp. 29-52.
- Fry, D. et al. "The Identification and Discrimination of Synthetic Vowels" (1962) Lg&Sp 5: 171-89.
- Furness, F. "Listening: A Case of Terminological Confusion" (1957) JEdPsy 48: 477-82.
- Garding, L. & Gerstman, L. "The Effect of Changes in the Location of an Intonation Peak on Sentence Stress" (1960) SLing 14: 57-9.
- Garrett, M. et al. "The Active Use of Grammar in Speech Perception" (1966) Perception and Psychophysics 1: 30-2.
- Glichsterg, D. A Study of the Span of Immediate Memory (1964) Ph.D. Diss. (Univ. of Michigan). In DissAb 24 (12): 520-41.
- Goldiamond, I. & Hawkins, W. "Voxierversuch: The Log Relationship Between Word-Frequency and Recognition Obtained in the Absence of Stimulus Words" (1958) JEXPsy 56 (6): 457-63.
- Grundstrom, A. "An Experimental Study of Interrogative Intonation in French" (1966) Ph.D. Diss. (Univ. of Michigan). In DissAb 28: 215-A.
- Farley, C. "Factorial Analysis of Speech Perception" (1956) JSHD 21: 76-87.
- Helmholtz, H. von Sensations of Tone (1912, 4th) London (Longmans, Green & Co.).
- Firsh, I. et al. "Intelligibility of Different Speech Materials" (1954) JASA 26: 530-8.
- Hofmann, Th. "Lexical Redundancy and the Perception of Speech" (1971) Proc. of 7th Congress of Phonetic Sciences, The Hague (Mouton & Co.).
- Jeffress, L. "Masking" (1970) in Tobias, pp. 87-114.
- Johnson, K. "The Effect of Classroom Training Upon Listening Comprehension" (1951) JComm 1 (1): 57-62.

- Johnson, N. "The Psychological Reality of Phrase-Structure Rules" (1966) JVIVB 4: 469-75.
- Ladefoged, P. "The Perceptor of Speech" (1959) Mechanization of Thought Processes, Vol. 1, London (HNSO) pp. 397-409.
- Ladefoged, P. Three Areas of Experimental Phonetics (1967) London (Oxford U. Press).
- Lane, H. "The Motor Theory of Speech Perception: A Critical Review" (1965) PsyB 72 (4): 275-309.
- Lewis, T. "Listening" (1958) RLDPes 28: 89-95.
- Lieberman, A. "Some Results of Research on Speech Perception" (1957) JASA 29 (1): 117-23.
- Lieberman, A. "Perception of the Speech Code" (1967) PsyR 74 (6): 431-61.
- Lieberman, P. "Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech" (1963) Lg&Sp 6: 172-87.
- Lieberman, P. Intonation, Perception and Language (1967) Cambridge (MIT Press / Research Mono. 38).
- Licklider, J. "On the Process of Speech Perception" (1952) JASA 24: 590-4.
- Lycns, J. & Wales, R. (Ed) Psycholinguistic Papers (Proc. of the 1966 Edinburgh Conference) (1966) Edinburgh (Edinburgh U.P.).
- Karlin, G. "A Factorial Study of Auditory Function" (1942) Psychometrika 7: 251-79.
- McMahan, L. "Grammatical Analysis as Part of Understanding a Sentence" (1963) Ph.D. Diss. (Harvard Univ.).
- Mackey, W. Language Teaching Analysis (1965) London (Longmans, Green & Co.).
- Falmer, B. "Analyse prosodique et analyse grammaticale" (1967) Word 23: 374-8.
- Miller, G. "The Masking of Speech" (1947) PsyB 44: 105-29.
- Miller, G. "The Magic Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information" (1956) PsyB 63: 81-97.
- Miller, G. "Some Psychological Studies of Grammar" (1962) AmPsy 17: 748-62.
- Miller, G. et al. "The Intelligibility of Speech as a Function of the Context of the Test Materials" (1951) JExpPsy 41: 329-35.
- Miller, G. & Isard, S. "Some Perceptual Consequences of Linguistic Rules" (1963) JVLVB 2: 217-28.
- Miller, G. & Nicely, P. "An Analysis of Perceptual Confusions among Some English Consonants" (1955) JASA 27: 338-52. (1961) Reprinted in Saporta pp. 153-75.
- Nash, R. "John Likes Mary More Than Bill. An Experiment in Disambiguation Using Synthesized Intonation Contours" (1970) Phonetica 22 (3): 170-88.
- Nichols, R. "Factors in Listening Comprehension" (1948) SpM 15: 154-63.
- Nichols, R. "Needed Research in Listening Communication" (1951) JComm 1 (1): 48-50.

- O'Connor, J. & Tccley, C. "The Perceptibility of Certain Word-Boundaries" (1964) in Abercrombie et al. pp. 171-6.
- Oldfield, R. & Marshall, J. (Ed) Language (1968) Hammondswoth (Perguin).
- Pickett, J. & Pcllack, I. "Intelligibility of Excerpts from Fluent Speech: Effects of Rate of Utterance and Duration of Excerpt" (1963) Lq&Sp 6: 151-64.
- Fimsleur, P. & Quinn, T. (Ed) The Psychology of Second Language Learning (1971) Cambridge (Cambridge U.P.).
- Eclitzer, R. & Weiss, L. "Developmental Aspects of Auditory Discrimination, Echo Response, and Recall" (1969) MLJ 53 (2): 75-85.
- Pcllack, I. & Pickett, J. "The Intelligibility of Excerpts from Conversation" (1963) Lq&Sp 6: 165-71.
- Pcllack, I. et al. "Intelligibility of Selected Message Sets" (1958) JASA 30: 643.
- Pcllack, I. et al. "Intelligibility of Known and Unknown Message Sets" (1959) JASA 31: 273-9.
- Rankin, P. "The Measurement of the Ability of Understanding Spoken Language" (1926) Ph.D. Diss. (Univ. of Michigan). In DissAb 12: 847-8.
- Rigault, A. "Réflexions sur le statut phonologique de l'intonation" (1964) Proc. 9th Cong. Ling., Cambridge, Mass. pp. 849-58.
- Fivers, W. "Listening Comprehension" (1966) MLJ 50 (4): 196-204.
- Fivers, W. "Linguistic and Psychological Factors in Speech Perception and their Implications for Teaching Materials" (1971) in Fimsleur & Quinn, pp. 123-34.
- Rosenzweig, M. & Ecstman, L. "Frequency of Usage and the Perception of Words" (1958) Science 127 (3293): 263-6.
- Rubenstein, H. & Pcllack, I. "Word Predictability and Intelligibility" (1963) JVLVB 2: 147-58.
- Rubenstein, H. et al. "Word Length and Intelligibility" (1959) Lq&Sp 2: 175-78.
- Saporta, S. (Ed) Psycholinguistics: A Book of Readings (1961) New York (Holt, Rinehart & Winston).
- Savin, H. "The Word Frequency Effect and Errors in the Perception of Speech" (1962) Ph.D. Diss. (Harvard Univ.).
- Spearritt, D. Listening Comprehension - A Factorial Analysis (1962) A.C.E.R. Research Series No. 76, Melbourne (G.W. Green & Sons).
- Spclsky, P. et al. "Preliminary Studies in Development of Techniques for Testing Overall Second Language Proficiency" (1968) Language Learning, Special Issue, No. 3: 79-101.
- Still, E. "The Relationship Between Listening Ability and High School Grades" (1955) Ph.D. Diss. (Univ. of Pittsburgh). In DissAb 15: 1761-2.

- Svensson, S-G. "A Preliminary Study of the Role of Prosodic Parameters in Speech Perception" (1971) Speech Transmitter Lab Quarterly Progress and Status Report (Dept. of Speech Communication, Stockholm, Sweden), 2-3, pp.24-42.
- Tchias, J. (Ed) Foundations of Modern Auditory Theory, Vol. 1 (1970) New York & London (Academic Press).
- Thorne, J. "On Hearing Sentences" (1966) in Lyons & Wales pp. 3-28.
- Treisman, A. "The Effect of Irrelevant Material on the Efficiency of Selective Listening" (1964) AmJpsy 77: 533-46. (1968) Reprinted in Oldfield & Marshall, J. pp. 132-46.
- von Békésy, G. "The Ear" (1957) ScAm 197 (2, Aug): 66-78.
- Warren, R. "Perceptual Restoration of Missing Speech Sounds" (1970) Science 167: 392-3.
- Warren, R. & Warriner, B. "Auditory Illusions and Confusions" (1970) ScAm 223 (6, Dec): 30-6.
- Wilkinson, A. & Atkinson, D. "A Test of Listening Comprehension" (1966) EdR 18: 177-85.
- Witting, C. "On the Auditory Phonetics of Connected Speech: Errors and Attitudes in Listening" (1962) Word 18: 221-48.

Journal Abbreviations

<u>AJPSY</u>	American Journal of Psychology
<u>AmPSY</u>	American Psychologist
<u>DissAb</u>	Dissertation Abstracts
<u>EdF</u>	Educational Review
<u>JASA</u>	Journal of the Acoustic Society of America
<u>JCCrn</u>	Journal of Communications
<u>JEdPSY</u>	Journal of Educational Psychology
<u>JExpSY</u>	Journal of Experimental Psychology
<u>JSEd</u>	Journal of Speech & Hearing Defects
<u>JVLVB</u>	Journal of Verbal Learning & Verbal Behaviour
<u>Lg&Sp</u>	Language & Speech
<u>MLJ</u>	Modern Language Journal
<u>PSYB</u>	Psychological Bulletin
<u>PSYR</u>	Psychological Review
<u>REdRes</u>	Review of Educational Research
<u>SLing</u>	Studia Linguistica
<u>ScAm</u>	Scientific American
<u>SPM</u>	Speech Monographs

APPENDICES

FEUILLE DE RESPONSES

S _____

1. a.

b.

c.

commentaire(s)

.....

2. a. _____

b. _____

c. _____

.....

3. a. _____

b. _____

c. _____

.....

4. a. _____

b. _____

c. _____

.....

5. a. _____

b. _____

c. _____

.....

6. a. _____

b. _____

c. _____

.....

7. a. _____

b. _____

c. _____

.....

8. a. _____

b. _____

c. _____

.....

9. a. _____

b. _____

c. _____

.....

FEUILLE DE REVISIONS

Discrimination auditive

- I. Exemples: a) B est plus bas que A
plus haut
- b) B est plus fort que A
plus faible

1. B est plus fort et plus haut que A.
plus faible plus bas

2. B est plus fort et plus haut que A.
plus faible plus bas

3. B est plus fort et plus haut que A.
plus faible plus bas

4. B est plus fort et plus haut que A.
plus faible plus bas

II.
1. B est plus fort et plus haut que A.
plus faible plus bas

2. B est plus fort et plus haut que A.
plus faible plus bas

3. B est plus fort et plus haut que A.
plus faible plus bas

4. B est plus fort et plus haut que A.
plus faible plus bas

III. A. _____ B. _____
1. _____ C. _____ D. _____
2. _____

IV. lit, thé, terre, la, dos, loup, lu, deux

1. La voyelle correspond à celle du mot _____

2. " " " " " " _____

3. " " " " " " _____

4. " " " " " " _____

5. La voyelle correspond à celle du mot _____

6. " " " " " " " _____

7. " " " " " " " _____

8. " " " " " " " _____

9. La voyelle correspond à celle du mot _____

10. " " " " " " " _____

V. a _____

b _____

VI. 1. a _____ b _____

2. a _____ b _____

3. a _____ b _____

4. a _____ b _____

5. a _____ b _____

6. a _____ b _____

7. a _____ b _____

8. a _____ b _____

9. a _____ b _____

10. a _____ b _____

11. a _____ b _____

12. a _____ b _____

13. a _____ b _____

14. a _____ b _____

15. a _____ b _____

Formulation d'hypothèses

- I. 1. _____ 2. _____
 3. _____ 4. _____
 5. _____

- II. 1. _____
 2. _____
 3. _____
 4. _____
 5. _____
 6. _____
 7. _____
 8. _____
 9. _____
 10. _____
 11. _____
 12. _____
 13. _____
 14. _____
 15. _____
 16. _____

III. Exemple: Je parle du nombre - Je parle d'une ombre

1. On s'en dégoûte - _____
 2. Il est ouvert - _____
 3. Cesse-là - _____
 4. Sa toilette est faite - _____
 5. Il l'armait - _____

6. Il parle d'une autre

- _____

7. Il ficela.

- _____

8. C'est un moi effort

- _____

9. Laisse-les rentrer

- _____

10. C'est un neuf

- _____

APPENDIX 2A LISTENING COMPREHENSION ANSWERS

1. Nous n'allons pas nous chamailler à cause d'eux.
2. C'est à Greux qu'est l'église principale.
3. Christ était émeutier.
4. Tu n'as pas eu le mal de mer?
5. Voici Alexandre au Mexique.
6. Ca te plaît moins que la perche?
7. C'est du bon lait normand.
8. Le rinçage est bien plus facile.
9. C'est le coup du chimpanzé.
10. Eh bien, pas si bête, mon cher.
11. Adressez-vous aux tribunaux.
12. Voici la fourchette de France-Inter.
13. Pour un parfumeur, quelle enseigne!
14. Je suis hérissée, il n'y a rien à faire!
15. Dans moins d'une heure, je saurai tout.
16. Les médecins nagent, mon vieux.
17. Comment ça s'écrit ton nom, déjà?
18. Où dessine encore Rembrandt.
19. Vous êtes bien du canton?
20. Comme un pot?
21. Un peu mieux que du côté droit.
22. On dirait que le train tout entier vous roule sur le corps,
23. J'aurais pas fait étalage de mon héroïsme.
24. Tu as du feu, toi?
25. Passez-moi la grenade!
26. Viens, papa!
27. Les vivants sont rares.
28. Alors, ceux-là aussi vous pèsent?
29. Pour les loges ou Pour l'éloge
30. Aide-moi à dégrafer mon boa.
31. Socrate était donc un chat?
32. Et quand vous êtes-vous cogné?
33. Vos veines ont l'air de se gonfler.
34. Qui: On?

35. Vous avez hérité les actions.
36. Sens-moi cette fraîcheur.
37. Qui remarquâtes-vous de particulier? (Qu'y)
38. Ah, des floppées!
39. Avec deux t. (thés)
40. Ah, c'est pas à nous que ces oublis-là arriveraient! Ah non alors.
41. Ah, le monstre, il me trompe!
42. Et un chat aliéné de toutes ses pattes?
43. N'exagérons rien.
44. Tâte mes mains.
45. En voilà la première nouvelle.

APPENDIX 2B AUDITORY DISCRIMINATION TEST ANSWERS

- I. 1. B est plus fort et plus bas que A.
 2. B est plus faible et plus haut que A.
 3. B est plus fort et plus haut que A.
 4. B est plus faible et plus bas que A.
- II. 1. B est plus faible et plus haut que A.
 2. B est plus fort et plus bas que A.
 3. B est plus faible et plus bas que A.
 4. B est plus fort et plus haut que A.
- III. 1. C
 2. D
- IV. 1. La voyelle correspond à celle du mot lu.
 2. La voyelle correspond à celle du mot lit.
 3. La voyelle correspond à celle du mot loup.
 4. La voyelle correspond à celle du mot dos.
 5. La voyelle correspond à celle du mot dos.
 6. La voyelle correspond à celle du mot thé.
 7. La voyelle correspond à celle du mot la.
 8. La voyelle correspond à celle du mot terre.
 9. La voyelle correspond à celle du mot deux.
 10. La voyelle correspond à celle du mot lu.
- VI. 1. sɔrɛ
 2. sɥi
 3. mɔ̃vjɔ
 4. kris
 5. satplɛ
 6. kɑ̃tɔ̃
 7. vurulsy
 8. ɔrɛpaf
 9. tyady
 10. vjɛ̃
 11. vɑ̃sɔ̃
 12. mɔ̃boa
 13. sʔkra
 14. kɔ̃pɛ
 15. tɛʃæɾ

APPENDIX 2C HYPOTHESIS FORMULATION TEST ANSWERS

II. 1. Est-ce que tout fond vite?

2. L'homme a nui à son air.

3. Tel fut l'hiver ibère.

4. Pourquoi t'es-tu tu?

5. Quel pou dur à tuer.

6. Convain

7. Quel hymne est-ce?

8. Quand la laitue naît, le ver sort.

9. Qui terre a, guerre a.

10. Etant mou, il échoue.

11. Jamais lard n'alla mieux aux pois.

12. Ce goût doux est pis que tout.

13. Mille ans de paix c'est trop.

14. Qui dit khaki dit sale.

15. Ecoutez l'avis air-terre.

16. Les en louer m'est dur.

III. 1. On s'en dégoûte.

On sent des gouttes.

2. Il est ouvert.

Il est tout vert.

3. Cesse-là.

C'est cela.

4. Sa toilette est faite

Sa toile était faite.

5. Il l'armait.

Il la remet.

6. Il parle d'une autre.

Il parle du nôtre.

7. Il ficela.

Il fit cela.

8. C'est un mol effort.

C'est un mollet fort.

9. Laisse-les rentrer.

Laisse l'air entrer.

10. C'est un neuf.

C'est un oeuf.

APPENDIX 3

BMD02D CORRELATION WITH TRANSGENERATION - REVISED JANUARY 29, 1970
HEALTH SCIENCES COMPUTING FACILITY, UCLA

PROBLEM CODE CORR

NUMBER OF VARIABLES 19

NUMBER OF CASES 14

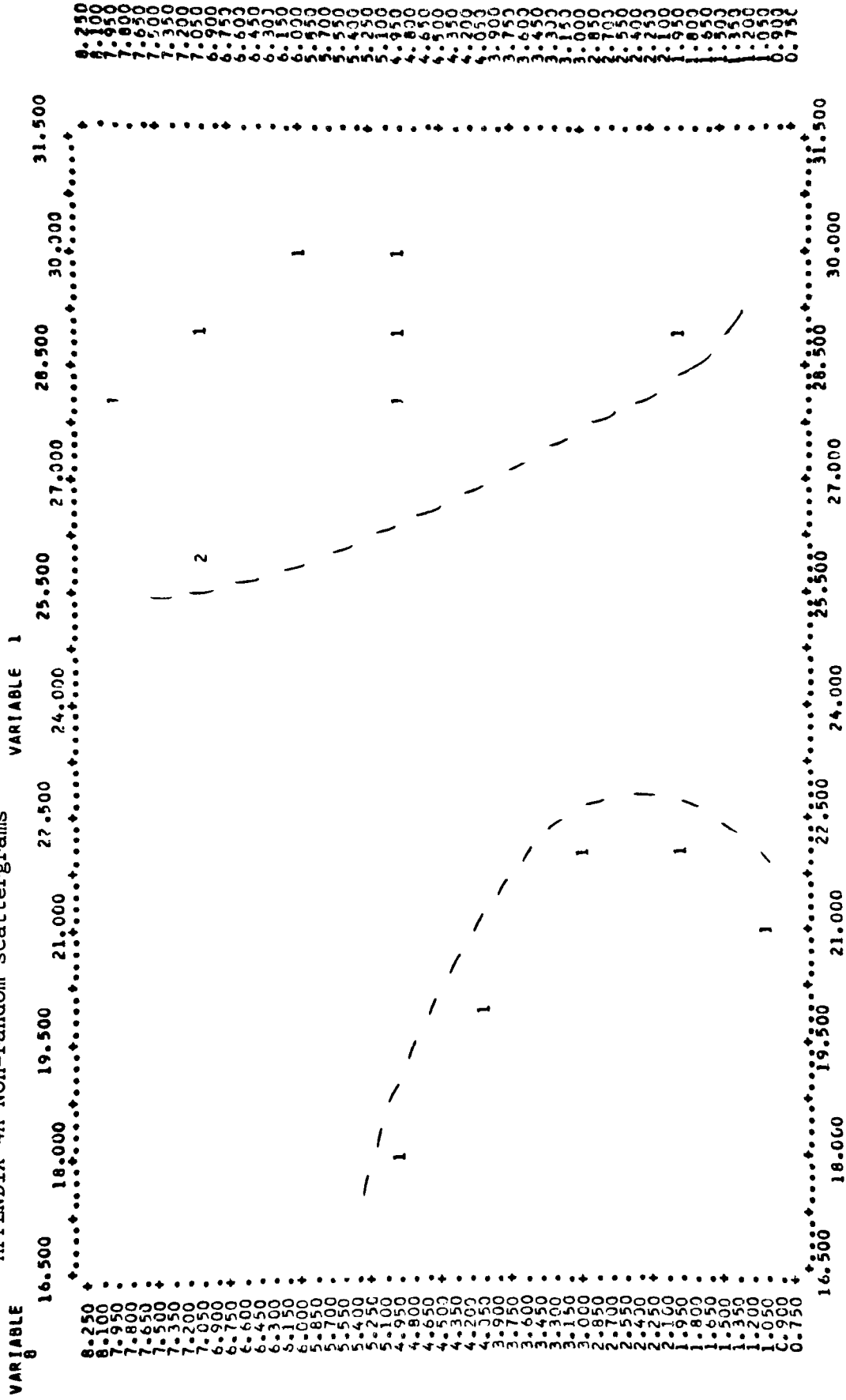
CORRELATION MATRIX

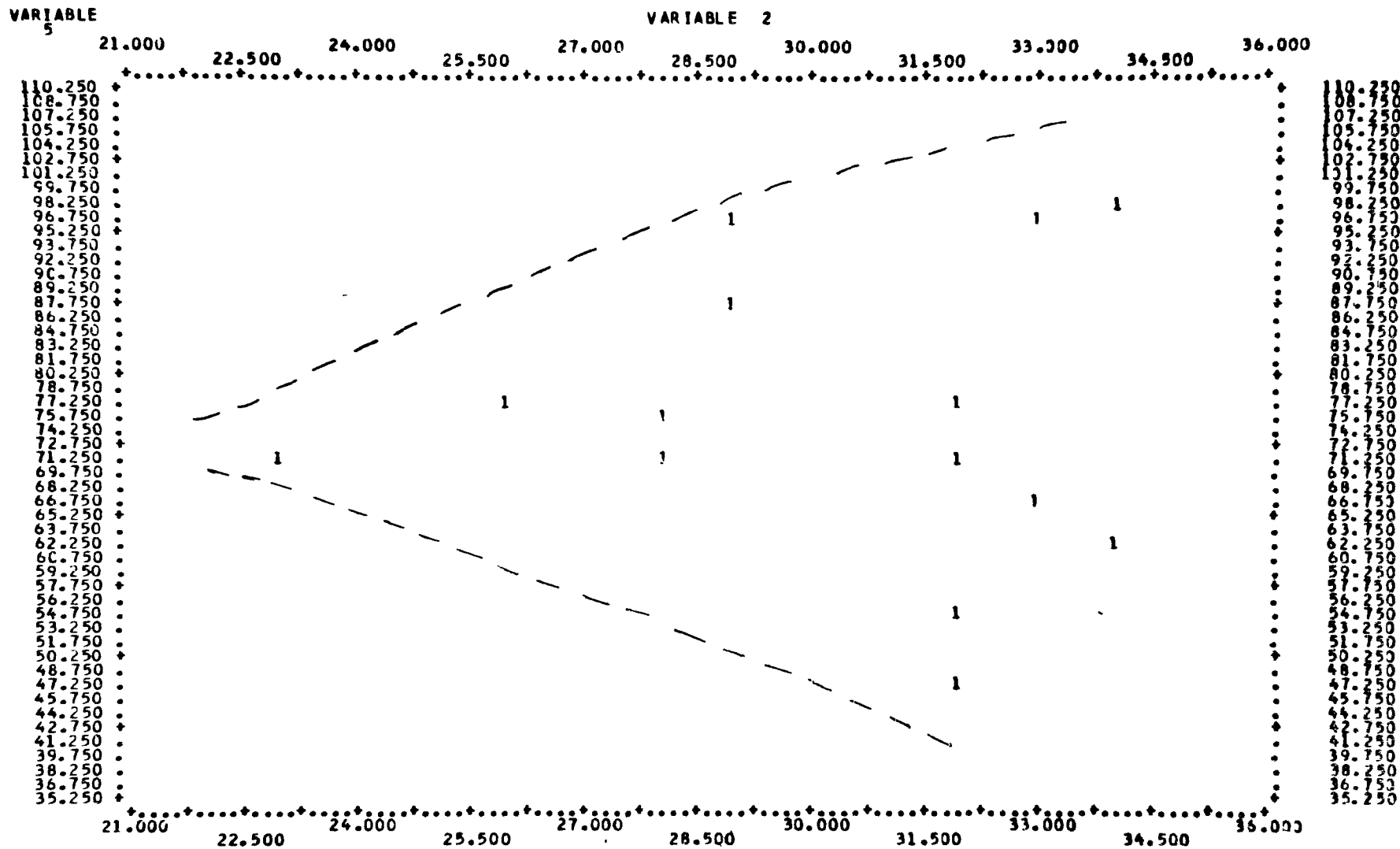
	V.1	V.2	V.3	V.4	V.5	V.6	V.7	V.8	V.9
V.1	1.0	.810	.824	.787	-.116	-.178	.230	.447	.501*
V.2		1.0	.900	.810	-.046	-.030	.274	.404	.620**
V.3			1.0	.953	.058	.038	.252	.577**	.741***
V.4				1.0	.056	-.050	.225	.659**	.759***
V.5					1.0	.893	.231	.154	-.149
V.6						1.0	.249	.034	-.051
V.7							1.0	.283	.135
V.8								1.0	.578**
	V.10	V.11	V.12	V.13	V.14	V.15	V.16	V.17	V.18
V.1	.155	-.209	-.370	.064	.403	-.161	.140	-.216	-.409
V.2	.239	.073	-.089	.294	.196	-.106	.184	-.463	-.469
V.3	.464	-.151	-.283	.020	.147	-.101	.001	-.406	-.451
V.4	.507	-.179	-.314	-.071	.095	-.135	-.053	-.321	-.381
V.5	-.100	.134	.089	-.207	-.371	.694***	-.168	.061	.098
V.6	.014	.291	.267	-.030	-.293	.743***	-.009	.095	.168
V.7	.065	.393	.371	.418	.181	.098	.388	.026	-.052
V.8	.440	-.031	-.179	-.107	.092	.147	.043	-.343	-.220
V.9	.723	.036	-.051	.105	-.215	-.135	.173	-.201	-.288
V.10	1.0	-.021	-.039	-.185	-.003	.030	-.016	-.345	-.422
V.11		1.0	.958	.810	.083	.474	.724	.246	.276
V.12			1.0	.820	.127	.381	.726	.342	.315
V.13				1.0	.432	.042	.838	.246	.159
V.14					1.0	-.038	.748***	.208	-.145
V.15						1.0	.289	.100	.181
V.16							1.0	.308	.084
V.17								1.0	.787***

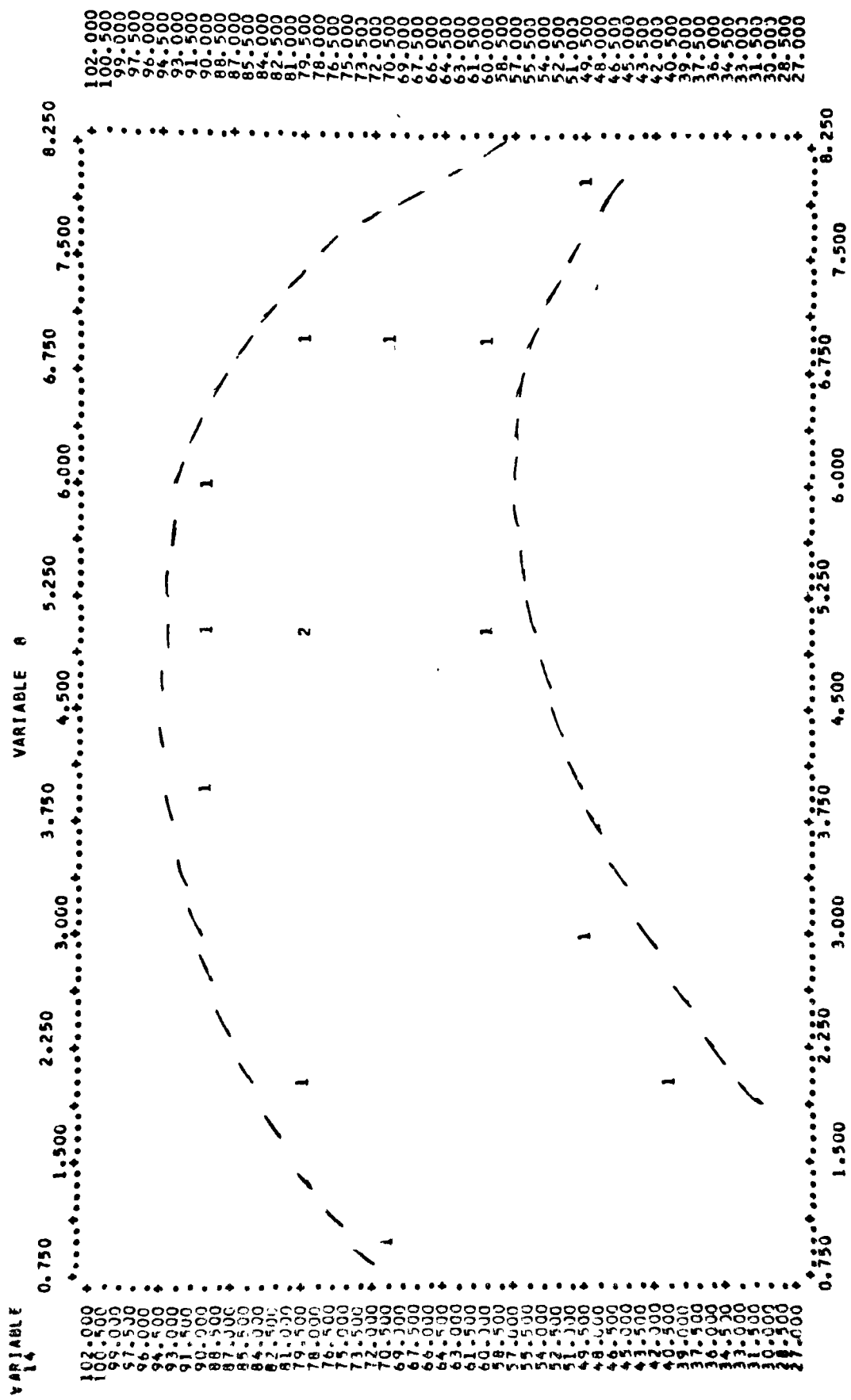
List of Variables

- V.1 Listening Comprehension Test Score, trial 1
- V.2 Listening Comprehension Test Score, trial 2
- V.3 Listening Comprehension Test Score, trial 3
- V.4 Listening Comprehension Test Score, 3 trials
- V.5 Amount of improvement in Listening Comprehension Test
between trials
- V.6 Amount of corrected improvement in Listening Comprehension
Test
- V.7 Auditory Discrimination Test, part 1 pure tones
- V.8 Auditory Discrimination Test, part 2, synthetic vowels
- V.9 Auditory Discrimination Test, part 3, segments, trial 1
- V.10 Auditory Discrimination Test, part 3, segments, trial 2
- V.11 Hypothesis Formulation Test, part 2
- V.12 Hypothesis Formulation Test, part 2
- V.13 Hypothesis Formulation Test, part 2
- V.14 Hypothesis Formulation Test, part 3, printed sentences
- V.15 Degree of imagination in Hypothesis Formulation Test
- V.16 Hypothesis Formulation Test, parts 2 & 3
- V.17 Age
- V.18 Number of years away from France
- V.19 Sex

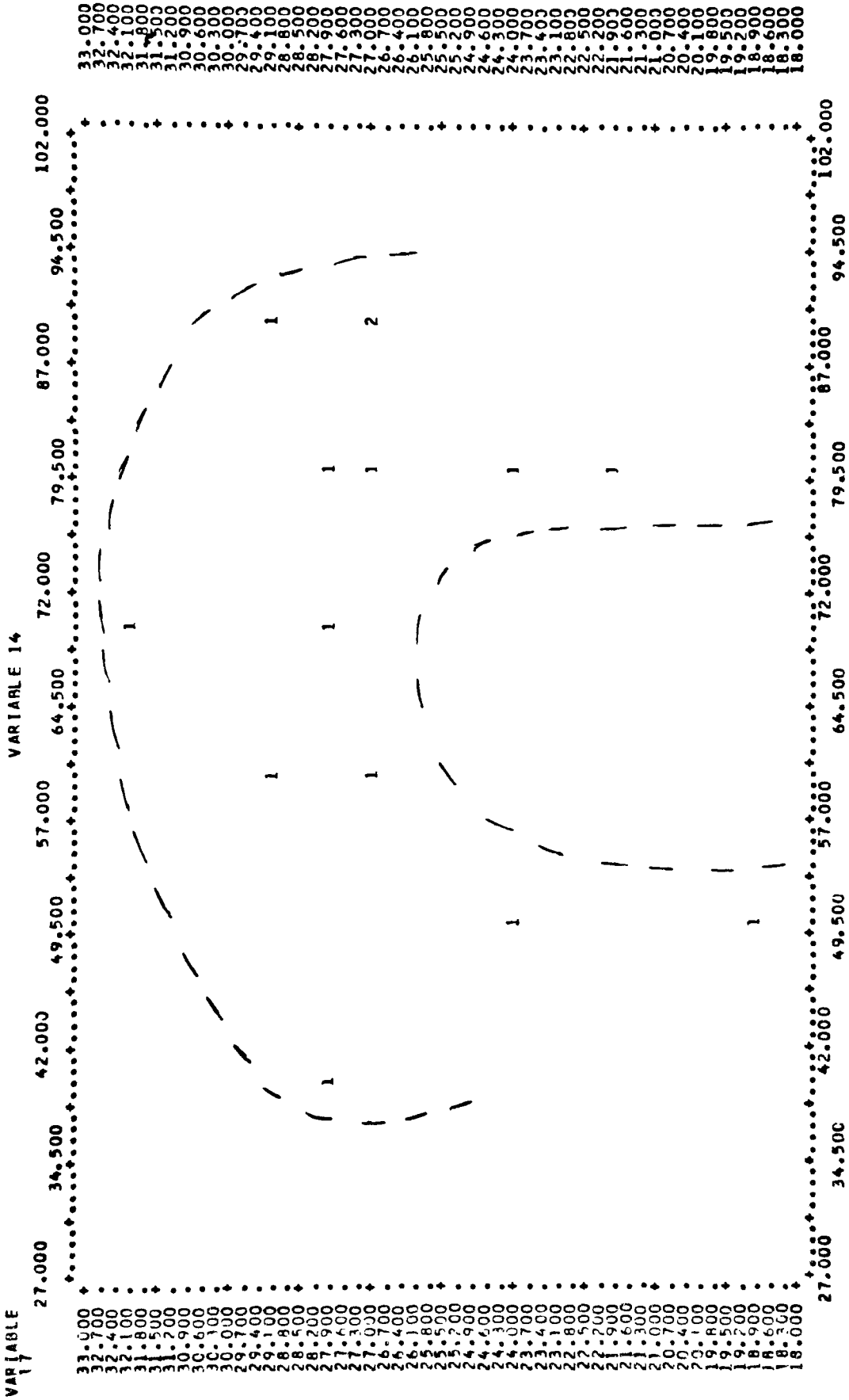
APPENDIX 4A Non-random scattergrams



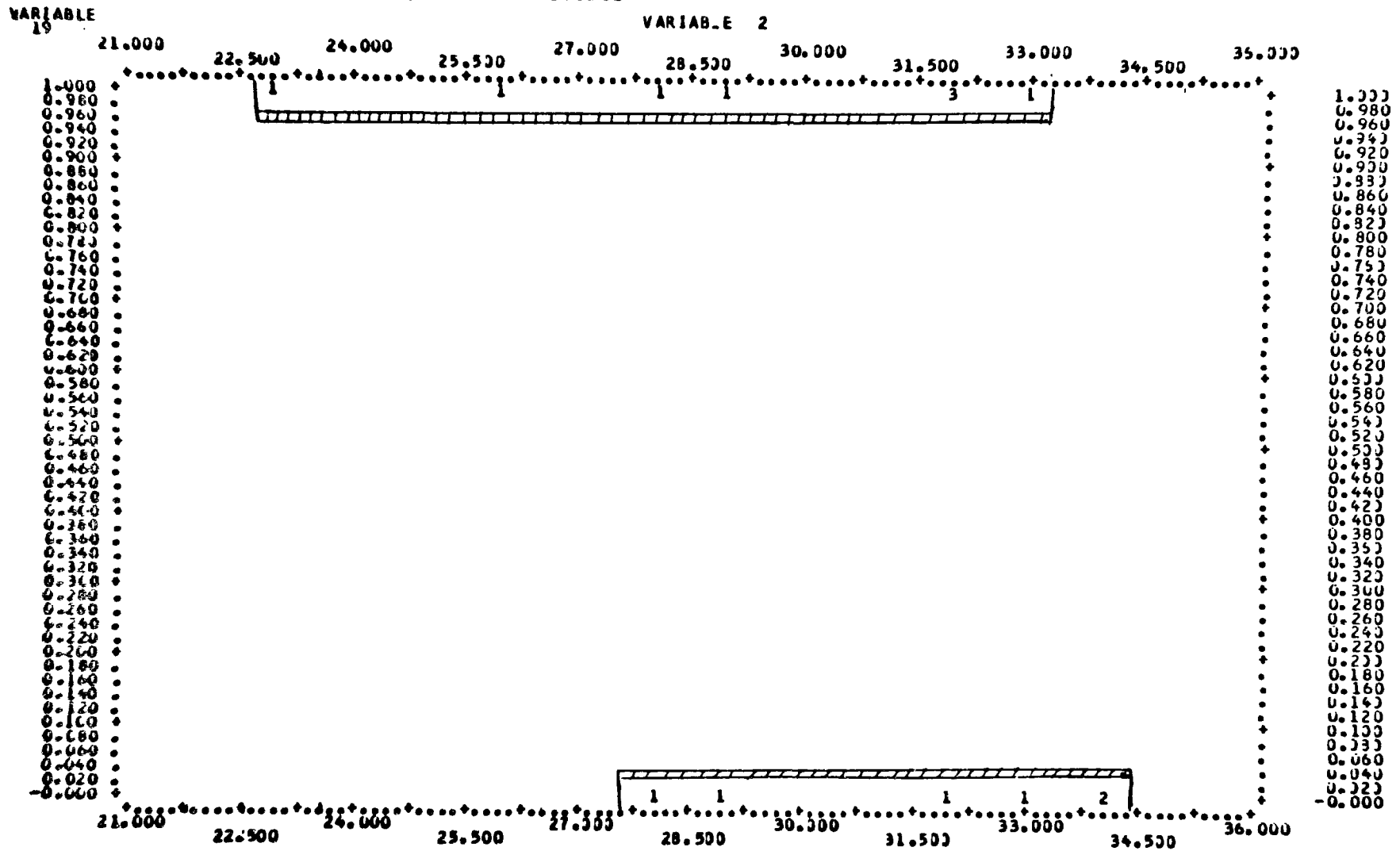


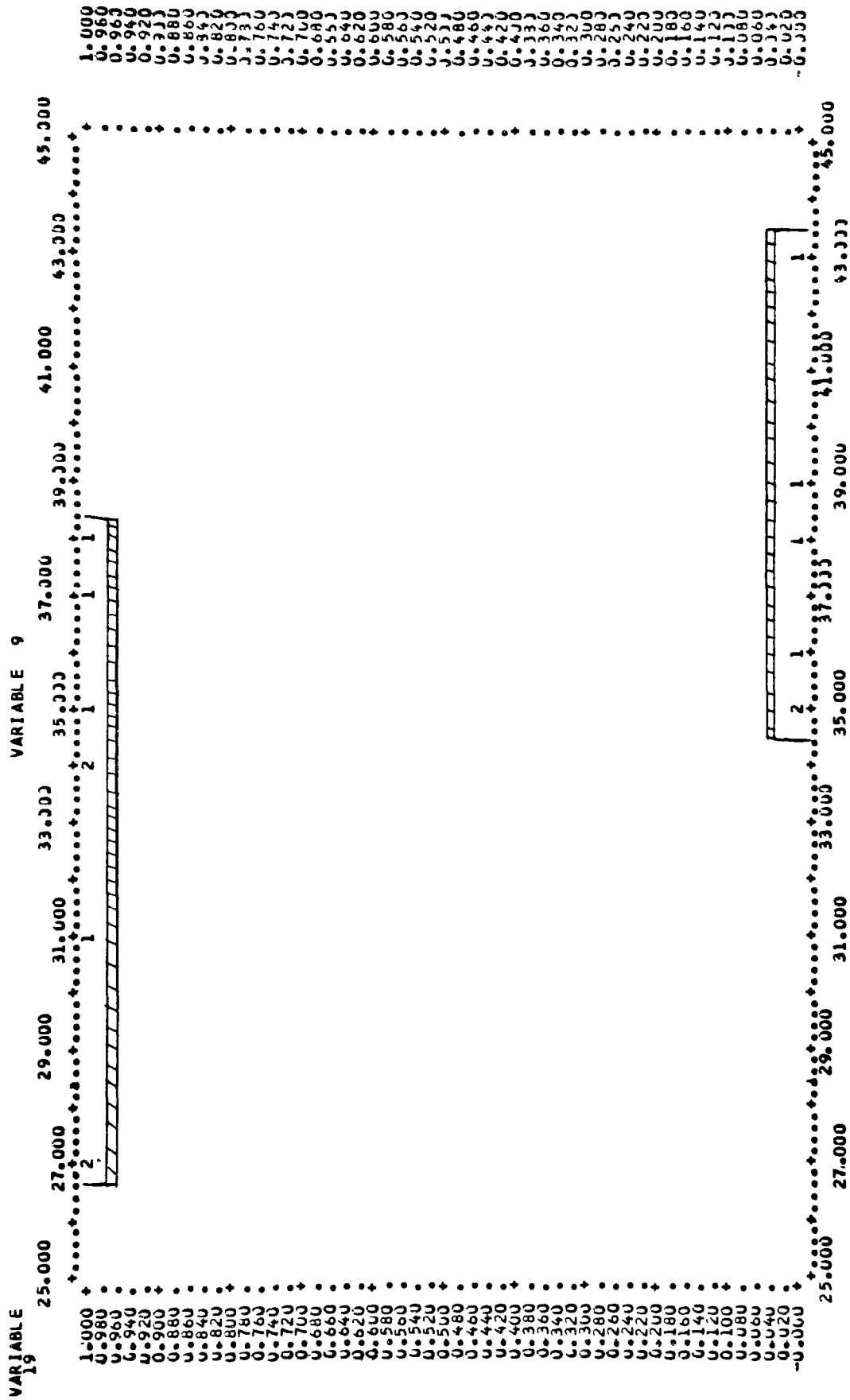


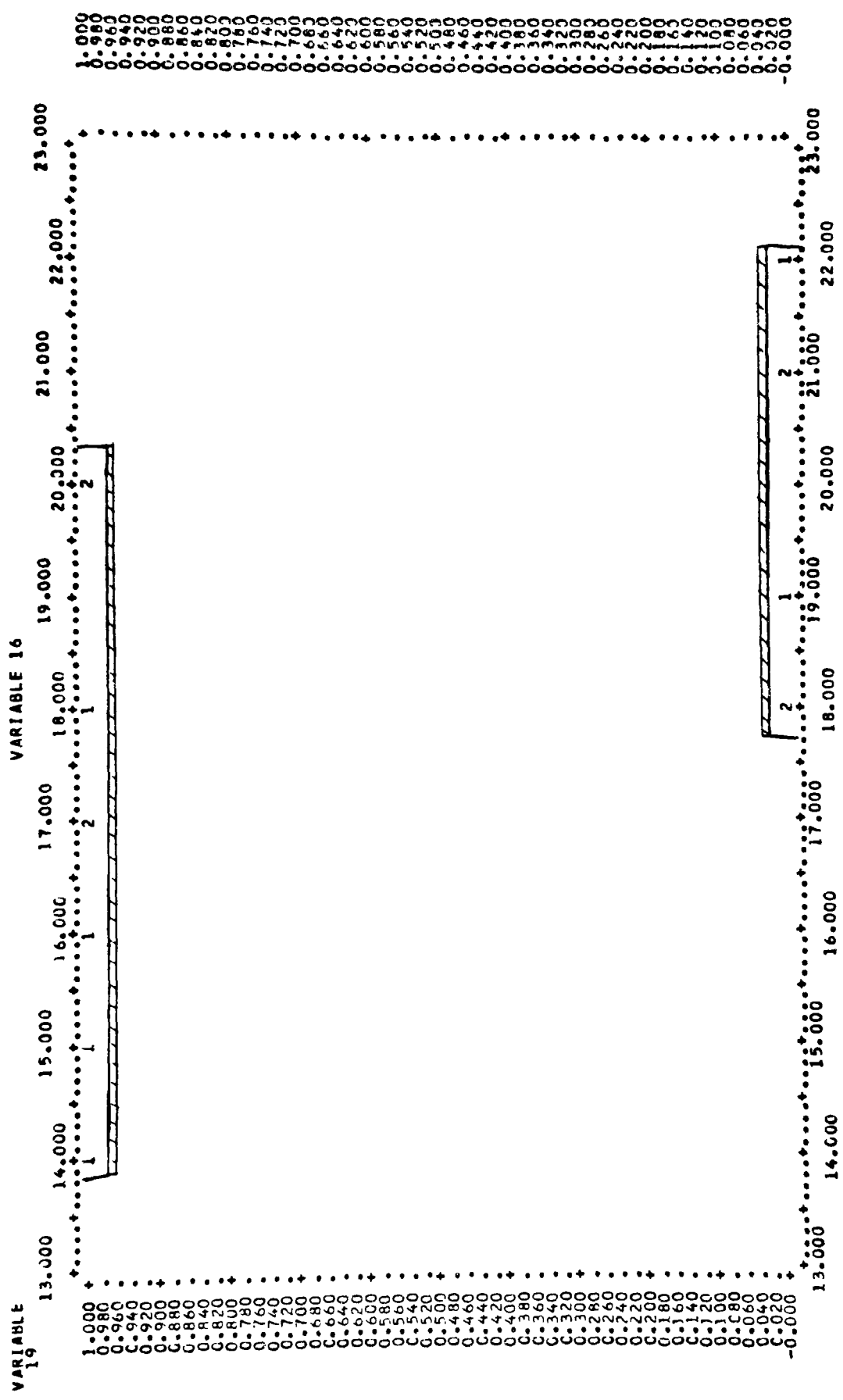
VARIABLE
14



APPENDIX 4B Sex (V.19) and test scores







ABSTRACT

Native speakers of European French were given tests of listening comprehension, auditory discrimination, and hypothesis formulation. Scores from these tests and personal data about the subjects were correlated to see what factors were related to listening comprehension of recorded material.

General tendencies noted were that auditory discrimination correlated with listening comprehension as measured by our tests, but that hypothesis formulation as defined by our testing procedures did not. However, subjects with low auditory discrimination scores did tend to have higher hypothesis formulation scores. One important conclusion was that so-called "native" listening comprehension ability could not be defined and therefore should not be used as a goal in foreign language methods. It was questioned whether ability to understand such difficult material as recorded films was a feasible goal in foreign language teaching.