63 3 3

BOLT BERANEK AND NEWMAN INC

CONSULTING · DEVELOPMENT · RESEARCH

402 604

AD NO.

RADC-TDR-63-90

AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES

TECHNICAL REPORT NO. RADC-TDR-63-90
25 February 1963

Prepared for:

Rome Air Development Center
Research and Technology Division
Air Force Systems Command
United States Air Force
Griffiss Air Force Base, New York

Project No. 4519, Task No. 45350

(Prepared under Contract No. AF 30(602)-2235)

CAMBRIDGE NEW YORK CHICAGO LOS ANGELES

NOTICES

Title of Report AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES

RADC-TDR-63-90

PUBLICATION REVIEW

This report has been reviewed and is approved.

Approved:

EDWIN J. LOWICKI, Acting Chief
Advanced Analysis Branch
Advanced Development Laboratory

Approved:

for HENRY HOFFMANN, Jr., Chief
Advanced Development Laboratory
Directorate of Communications

FOR THE COMMANDER:

IRVING J. GABELMAN
Director of Advanced Studies

RADC-TDR-63-90

AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES

TECHNICAL REPORT NO. RADC-TDR-63-90
25 February 1963

Rome Air Development Center
Research and Technology Division
Air Force Systems Command
United States Air Force
Griffiss Air Force Base, New York

Project No. 4519, Task No. 45350

(Prepared under Contract No. AF 30(602)-2235
by Karl D. Kryter and Jay H. Ball of Bolt Beranek
and Newman Inc., Cambridge, Massachusetts)

AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES

ABSTRACT

The results of previous technical reports, prepared under this contract, are summarized and the results of tests of various recently developed speech compression systems are presented and analyzed.

The results obtained from an analysis of published data and theory on techniques for the compression of the bandwidth of speech and new data collected on this matter during the course of this contract can be summarized as follows:

a. Semi-vocoders, operating at 9600 bits/sec, and channel vocoders, at 2400 bits/sec, will provide speech of adequate intelligibility and quality for most military communications. The voice quality of the semi-vocoders will usually be somewhat superior to that of the channel vocoders.

b. Results with as yet incomplete systems suggest that the Tarasoff-Daguet technique and the narrow-band spectrum sampling technique will ultimately provide speech intelligibility and quality comparable to that of the semi-vocoder. These systems will also require about 9600 bits/sec channel capacity during transmission.

c. Formant-tracking vocoders operating at about 1200 bits/sec can probably be developed to the point where they will provide speech intelligibility comparable to that from the channel vocoders operating at 2400 bits/sec. At the present time formant-tracking vocoders operating at 1000 bits/sec do not provide for adequate speech intelligibility.

d. Transmitting a restricted set of speech "patterns," from a larger set obtained from a digitized vocoder-type of speech analyzer-synthesizer, may provide a reduction in the information transmission rate normally required by that particular vocoder. Although this "pattern" matching technique is still in its early experimental stages it appears that the normal bit rate for a given vocoder system may, by this technique, be reduced to 2/3 and possibly to 1/2 its normal magnitude.

TABLE OF CONTENTS

1. INTRODUCTION

The aims set forth for Contract USAF 30(602)-2235, "An Evaluation of Speech Compression Techniques," were:

(1) to determine the relative strength and weakness of presently available speech compression techniques,

(2) to evaluate these techniques as to possible future potential and expansion,

(3) to determine the best method for equipment development in the near future, and

(4) to determine the best areas for future intensive research effort.

The results of major efforts to meet these above goals have been previously reported:

(1) K. N. Stevens, "Review of Existing Speech Compression Systems," RADC-TN-60-197, October 1960, and

(2) K. N. Stevens, M. H. L. Hecker and K. D. Kryter, "An Evaluation of Speech Compression Systems," RADC-TDR-62-171, 1 March 1962.

These two reports present in detail a critical analysis of the state-of-art to date in both engineering techniques and theory for reducing the frequency bandwidth normally required to transmit intelligible speech. The aforementioned reports also discuss the best probable paths to follow in future research and development in this area.

The results of a comprehensive speech testing program performed under this contract on various speech compression systems were a major aid to our evaluation of speech compression techniques. This speech testing was conducted in a way that permitted direct and accurate intercomparisons among various speech compression techniques with respect to PB word intelligibility, talker recognition, general speech quality and last, but not least, confusions made amongst so-called nonsense syllables. These latter tests provided interesting "diagnostic" information about the weaknesses and strengths of the various speech compression systems.

Subsequent to the completion of RADC-TDR-62-171 in March 1962, a survey was made of U. S. Air Force research and development projects that might provide additional speech compression equipment that would be completed by 1 January 1963. It was proposed, as a follow-up to the work already accomplished, that these new speech compression systems be tested with PB intelligibility tests in a manner that would permit comparison of the results with those reported in TDR-62-171. It was also proposed that during this period (March 1962 to January 1963) the contractor would explore the use of some "peak-picking (formant-tracking) technique" in conjunction with the so-called narrow-band spectrum sampling system. The termination date of the subject contract was extended to March 1963 so that these proposed tasks could be prosecuted.

In the report to follow we will: (1) briefly summarize a few selected highlights of the previous technical notes and reports issued under this contract; (2) give the test results for speech bandwidth compression systems that have been developed in the past six months or so and which we were able to test in December 1962 and January 1963; and (3) present a description of the efforts made to combine a peak-picking/formant-tracking system with the spectrum sampling (narrow-band) system.

2. SUMMARY OF PREVIOUS TECHNICAL REPORTS

2.1 RADC-TN-60-197, Review of Existing Speech Compression Systems

In this report, K. N. Stevens summarizes in a chart (see Fig. 1) his analysis of speech intelligibility test results of various speech compression systems obtained up to 1960. This chart relates the channel capacity required for speech transmission in bits-per-second to percent PB word intelligibility scores.

Although PB word scores cannot be used in an "absolute" sense inasmuch as they are strongly influenced by the ability of a particular group of listeners and talkers, as well as by the amount of training received by the listeners on a particular system, the reader may be interested in comparing the predictions made by Stevens on the basis of previous results from the literature and the results obtained in our studies as reported in TDR-62-171 and Section 3 of this report. It might be noted that in no case did we find in our tests a system that exceeded the estimates plotted by Stevens in Fig. 1. In general, however, the test results agreed very well with the "predictions" and extrapolations given in Fig. 1.

2.2 RADC-TDR-62-171, An Evaluation of Speech Compression Systems

TDR-62-171 is a lengthy document, 115 pages plus appendices. The following speech compression systems were tested in the course of the studies presented in TDR-62-171:

(1) Reference (low-pass) system. This system consists of a Spencer-Kennedy Model 302 electronic filter, set for 1500 cps low-pass operation with a characteristic slope of -36 dB/octave.

(2) Channel vocoder. Two systems were tested: Model HY-2, Philco Company, courtesy of the U. S. National Security Agency, and Model HC-135, Hughes Aircraft Company, Communications Division. Each vocoder has a 2400 bits/sec digital output and an estimated 400 cps analog bandwidth.

(3) Semi-vocoder. General Dynamics Corporation, Stromberg-Carlson Division. This "base-band" vocoder has an estimated analog bandwidth of 900 cps.

(4) Formant vocoder. Melpar, Inc. The formant-tracking vocoder tested produced a digital information stream of 1000 bits/sec; the bandwidth for analog operation is approximately 140 cps.

(5) Spectrum sampling (narrow-band) system. Bolt Beranek and Newman Inc. The analog bandwidth utilization is 800 cps.

(6) Tarasoff-Daguet system. Courtesy of U. S. Army Signal Research and Development Agency. The estimated analog bandwidth is approximately 1000 cps.

The speech compression systems listed above were subjected to several types of tests. These tests were designed to measure:

(1) the intelligibility of phonetically balanced (PB) words,

(2) the intelligibility of nonsense syllables, with emphasis on the confusions made,

(3) the general quality of the processed signal,

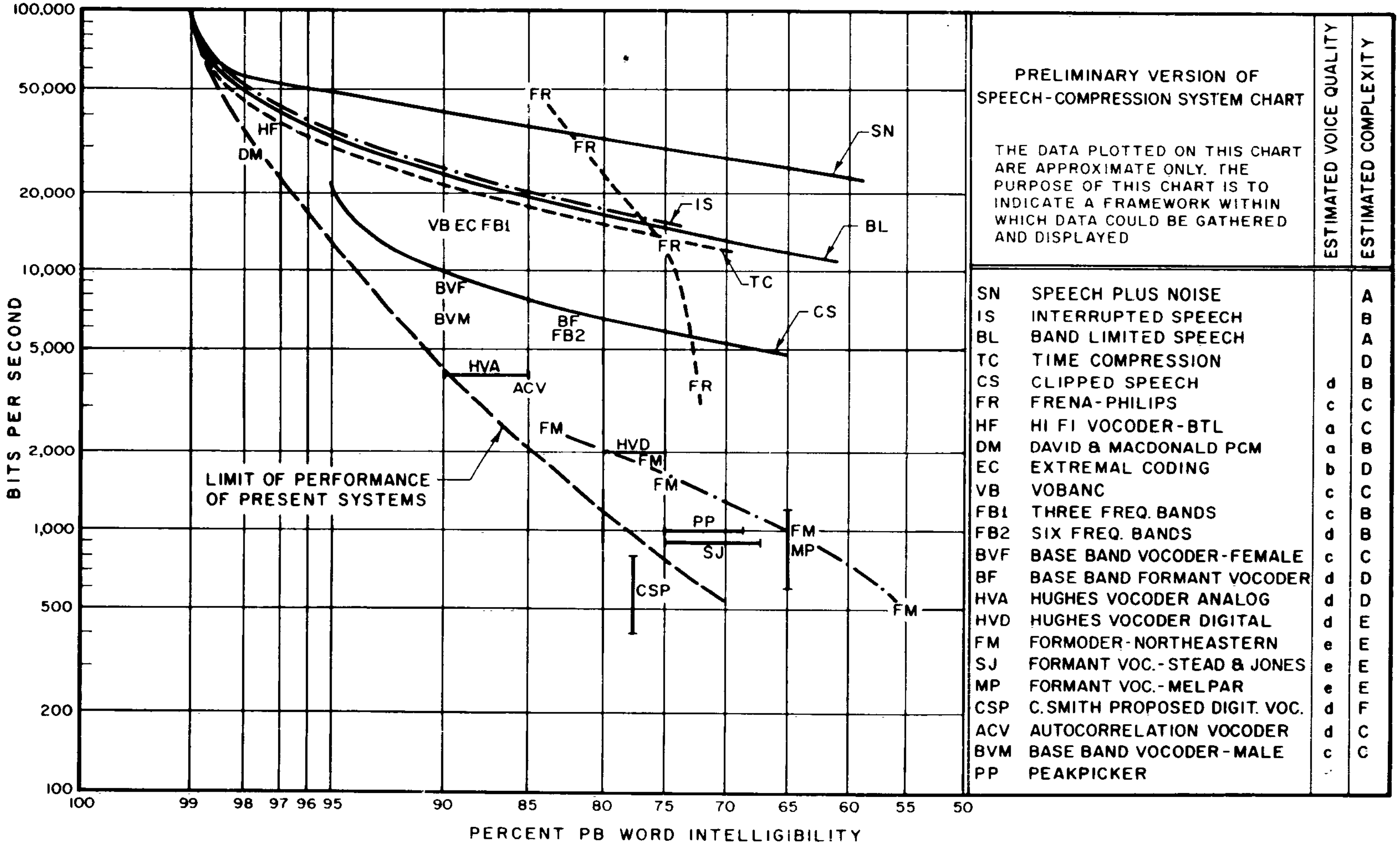

FIG. I ESTMIATED WORD INTELLIGIBILITY, VOICE QUALITY, AND EQUIPMENT COMPLEXITY OF VARIOUS SPEECH BANDWIDTH COMPRESSION SYSTEMS

(4) the accuracy with which listeners can recognize a given talker out of a small group of talkers, and

(5) the comprehension of continuous speech as a function of the degree of noise interference.

Figure 2, taken from TDR-62-171, shows the results of average scores obtained on PB word tests for male talkers on the various speech compression systems. This figure is represented as an aid in making comparisons with the speech compression system chart presented in Fig. 1 and with the test results given in Section 3 of this report.

There is much additional information available in TDR-62-171 of interest to research and engineering personnel working in the field of bandwidth compression systems. For our present summary we believe that the data presented in Fig. 2 will suffice. It can be noted, however, that the general rank ordering to be found in TDR-62-171 of the various systems by the different types of speech tests (talker recognition, general quality, etc.) were not appreciably different from the rank ordering found for these systems with the PB word tests.

It was concluded from these tests that channel vocoders operating at 2400 bits/sec and semi-vocoders at about 9600 bits/sec would provide adequate intelligibility and quality for most military communications but that formant-tracking vocoders (operated at about 1000 bits/sec) require considerable improvement before they can be considered satisfactory.

3. PB WORD TESTS OF "NEW" SYSTEMS

A survey of Air Force contractors working on speech communication systems revealed that the following speech compression devices being developed in this country were ready for testing by January 1963:

(1) Semi-vocoder: General Dynamics Corporation, Stromberg-Carlson Division.

(2) Semi-vocoder: Texas Instrument Company.

(3) Channel vocoder: Texas Instrument Company.

(4) Channel vocoder: General Dynamics Corporation, Stromberg-Carlson Division.

(5) Channel vocoder: Air Force Cambridge Research Laboratory (AFCRL).

(6) Pattern matching system: AFCRL

Accordingly, a contractor representative visited the above-mentioned organizations and obtained recordings of pre-recorded PB word tests as they were played through the various speech compression devices. The resulting recordings were then administered via Telephonics TDH-39 earphones to a crew of 6 trained listeners (college students) in January 1963. The tests were conducted under quiet listening conditions.

In addition to the test recordings obtained with the aforementioned systems additional PB word tests were administered to the listening crew over the so-called reference system and the Philco HY-2 channel-vocoder (see under Section 2 for descriptions of these systems). The tests with these latter two systems were included to determine whether this new

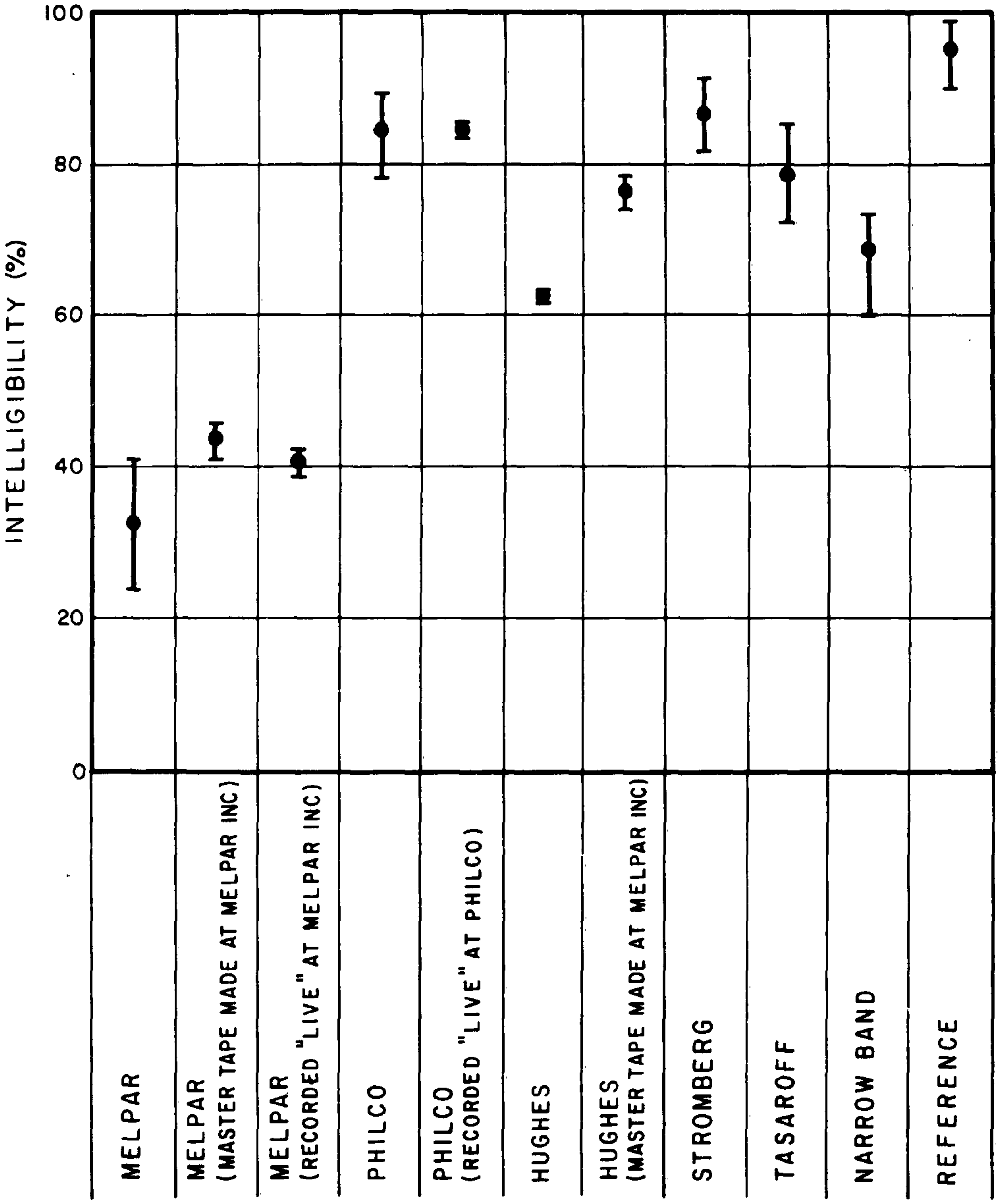

FIG. 2 AVERAGE SCORES ON PB WORD TESTS. MALE TALKERS IN QUIET. DYNAMIC MICROPHONE.

listening crew was more or less proficient than the crew used for the tests reported in TDR-62-171. The test results are presented in Figs. 3 and 4.

3.1 Discussion of Results

Statistical significance of differences. The grand averages given as the circles in Figs. 3 and 4 represent averages of four 50-word PB tests. Statistical analysis of PB word tests show that data points based on 200 PB words that differ by 5% points are significantly different at the 99% level of confidence; that is, a difference of 5 percentage points or greater between two systems or test conditions would be expected 99 times in a hundred repetitions of the same experiment. This general rule can be applied in estimating the possible significance of average differences among the results obtained with the various systems.

Some of the systems tested previously, with results reported in TDR-62-171, were re-tested along with these newer systems. The previous results, as well as the present ones, are shown on Figs. 3 and 4. The average difference between the grand averages for the systems that were tested in both experiments is 3.6 percentage points, with the difference in favor of the first study. If one wished to compare one of the new systems with one of those tested previously, but not re-tested in the present study, it would be reasonable to add about 3.6 percentage points to these new scores to equate the results for an apparent difference in test crew proficiency. After making such an adjustment, the statistic (5% points) suggested in the previous paragraph can be applied to the remaining difference, if any, to test for statistical significance.

<u>Absolute scores</u>. During the course of our experimentation with various speech compression systems it became obvious that with sufficient practice a crew of listeners could become quite proficient in interpreting the speech coming from a given system. For this reason, it was necessary to use an experimental design that provided the test crew with about an equal amount of listening experience for each of the bandwidth compression systems being evaluated. (The crew was given an initial 20 hours of training on the PB word tests with a broad-band speech system prior to the experiments proper to overcome the initial learning of the test materials and procedures usually encountered with a test crew of listeners.)

As a result of the experimental procedures followed, the scores we obtained of any given speech compression system are lower than the scores that might be obtained for that system with continued concentrated training of a group of listeners on only that or similar speech processing systems. However, our results provide, we believe, a valid basis for making comparisons among the various speech compression systems that were tested.

It is our subjective judgment that any system that would not score about 80% correct on PB word tests with a crew of listeners (such as used in these experiments) who were reasonably well trained on PB word tests but who heretofore had had no appreciable experience with a particular speech processing technique or system would be considered unacceptable by the average user. The "acceptability" of a speech communication system is, however, a complex question that cannot be answered on the basis of intelligibility tests alone, nor in any absolute sense. What is acceptable in some situations and by some users would be rated as unsatisfactory in other situations or by other users.

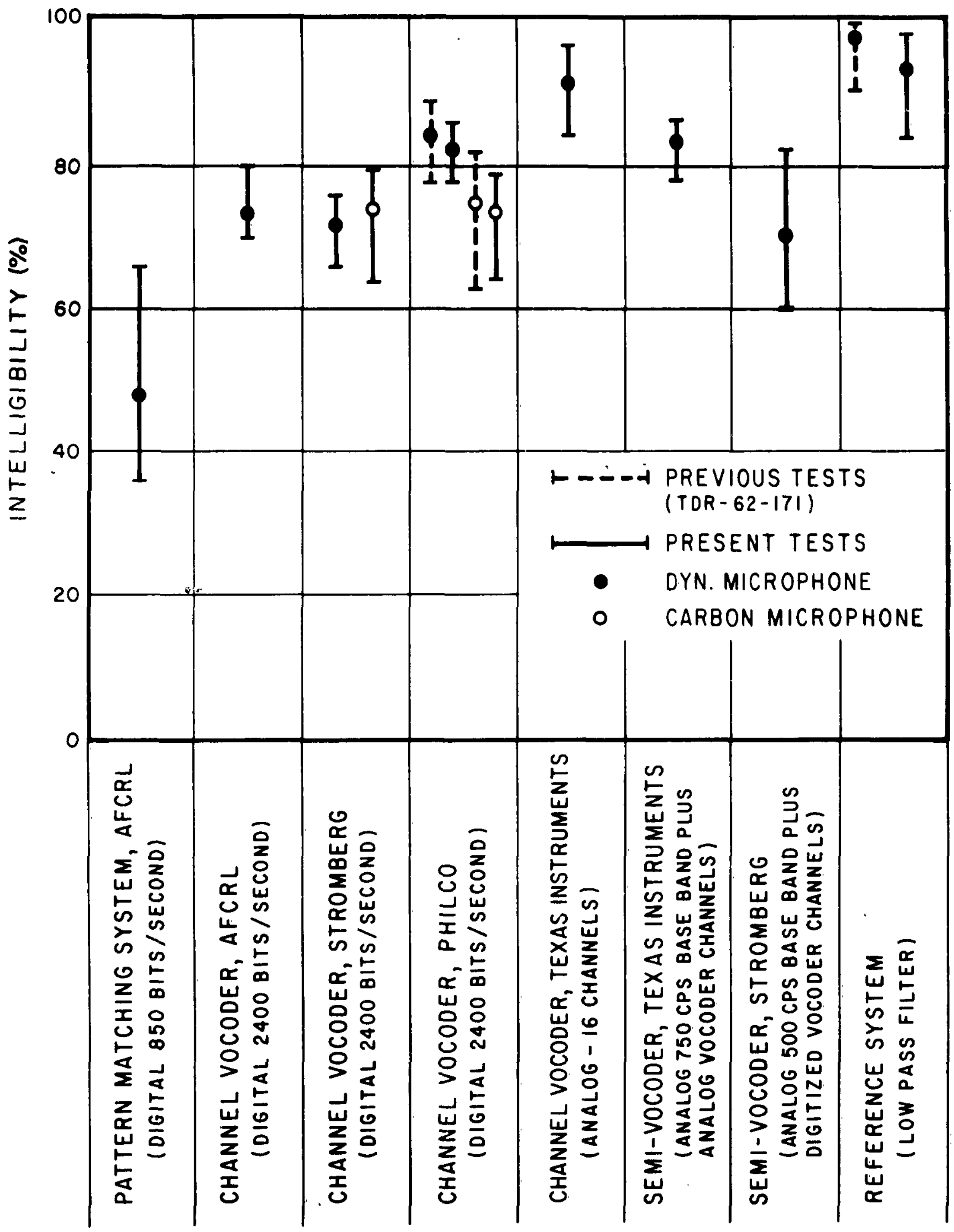

FIG. 3 AVERAGE SCORES ON PB WORD TESTS. TWO 50 WORD TESTS BY EACH OF TWO MALE TALKERS IN QUIET.

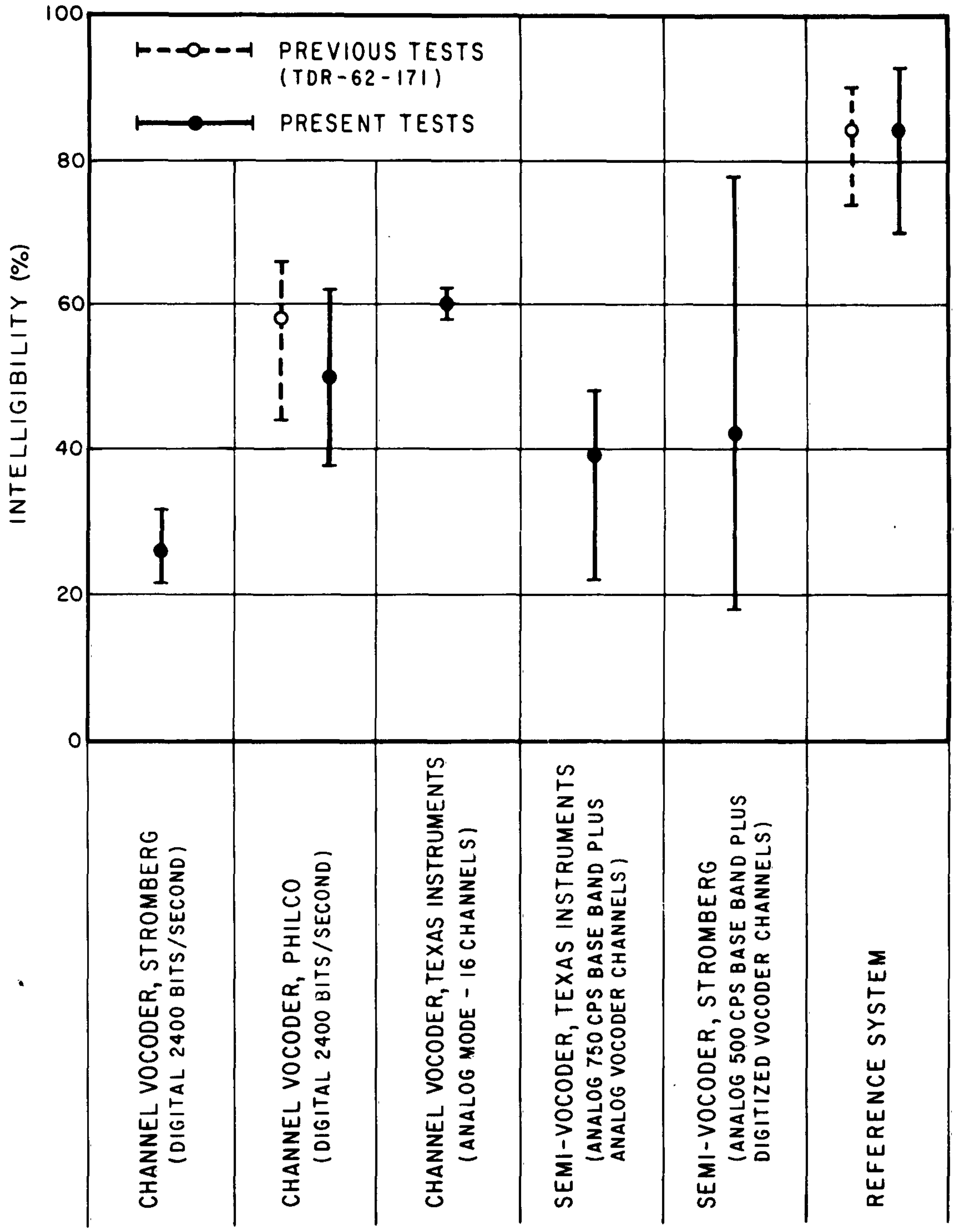

FIG. 4 AVERAGE SCORES ON PB WORD TESTS. TWO 50 WORD TESTS BY EACH OF TWO FEMALE TALKERS IN QUIET. DYNAMIC MICROPHONE.

3.1.1 Semi-vocoders -- male talkers

The results obtained with the semi-vocoders are worthy of several comments.

The semi-vocoder built by Stromberg-Carlson scored appreciably lower in these new tests than in the previous ones, 85% vs 70% (see Figs. 2 and 3). Presumably the major difference between the two tests lies in the fact that this semi-vocoder when first tested (Fig. 2) was operated in a completely analog mode whereas for the second tests (Fig. 3) the information from the vocoder channels was digitized during its processing by the system.

It will be recalled that a similar degradation resulting from switching from analog to digital mode was found in the Hughes vocoder (see TDR-62-171). Although we have no other comparative test scores to report that provide a direct evaluation of the effects of analog-to-digital conversion, it is apparent from these examples as well as other more casual tests we have had the opportunity to make on speech compression systems that the number of bits required to transmit in digital form the analog information involved is often greater than anticipated.

For this reason, intelligibility test scores obtained with systems that are similar, except that one is operated in the digital mode and the other in the analog mode, should not be compared. Therefore, the channel vocoder built by Texas Instrument Company cannot be judged as being superior to the other channel vocoders tested, although the scores were generally higher. Unfortunately, the analog-to-digital conversion equipment for the Texas Instrument Company channel vocoder was not completed at the time we had to conduct these tests, January 1963.

It was observed that although the voice quality of the semi-vocoder made by Texas Instrument Company was excellent the higher frequency bands sounded relatively weak. We think that it was for this reason that this semi-vocoder scored slightly lower than did the channel vocoder also made by The Texas Instrument Company. The higher frequency bands were relatively louder in this channel vocoder than in their semi-vocoder.

3.1.2 Channel vocoders -- male talkers

The channel vocoder made by Stromberg-Carlson Division of General Dynamics Corporation was designed for use with a standard carbon (telephone) microphone, whereas the Philco instrument was designed for use with a microphone, such as the Altec-Lansing 661A dynamic, that has a good frequency response down to frequencies as low as 50 cps. It is interesting to note that the Stromberg-Carlson channel vocoder performs about as well, if not slightly better, 74% vs 71%, with the carbon as it does with the dynamic microphone; on the other hand, the Philco channel vocoder scores 83% (average of 84% and 82%, see Fig. 3) with the dynamic microphone and 73% (average of 75% and 71%) with the carbon microphone. It should be noted, as might be expected, that with a carbon microphone the speech was more realistic (although the intelligibility was about equal) on the Stromberg-Carlson machine than on the Philco, but with a dynamic microphone the Philco channel vocoder was superior.

3.1.3 Pattern matching system

C. P. Smith* designed and developed a system in which the most frequent patterns (taken every fraction of a second) to be found during normal

---

* C. P. Smith, Speech data reduction: voice communications by means of binary signals at rates under 1000 bits per second. U. S. Air Force Cambridge Research Center, TR-57-111, ASTIA Doc. AD117290, Jan. 1957; and C. P. Smith, a method for speech data processing by means of a digital computer. Report ERD-TM-58-103, U. S. Air Force Cambridge Research Center, 1958.

speech at the output of a digitized channel vocoder are cataloged and stored in a magnetic memory device. The number of patterns stored can be limited to any desired number.

Subsequent to this initial cataloging and storage, the patterns forthcoming from a digital vocoder processing speech are compared with the patterns stored in the memory of a computer. The code for those patterns in the memory that come closest to matching those of the incoming signal from the vocoder is transmitted to the receiver portion of the system -- a second channel vocoder which resynthesizes the speech.

For these tests C. P. Smith played five PB word test tapes (Nos. 1E through 5E, 1 male talker) through his pattern analysis and storage device. He limited the stored information for spectral patterns to 600 bits and that for pitch to 250 bits. Following this compilation of most frequently occurring patterns, nine PB word tests (five of the same used for making the pattern analysis and catalog) were played through the entire system -- vocoder analysis-pattern matching-vocoder resynthesis. The output of the total system was recorded. In addition four of the PB tests were played through the channel vocoders operated "back-to-back" at 2400 bits/sec.

The scores obtained with these recordings are given in Table 1. It is clear from the results in Table 1 that the pattern matching technique, as presently operating, degrades PB word scores by a significant amount (average of 73% for the channel vocoder without the pattern matching system vs 48% with). Attention is invited to the fact that list 1E scored best on both systems.

Table 1

Word Scores AFCRL Speech Compression Systems

AFCRL Channel Vocoder plus Pattern Matching System 850 bits/sec
(250 bits pitch, 600 bits channel spectra)

PB List No.	% Correct	
1E	66	
2E	56	
3E	50	
4E	43	
5E*	36	Average 51%

*List 1 through 5E used for compiling vocoder patterns stored in memory of computer

PB List No.	% Correct	
6E	52	
7E	50	
8E	40	
9E	38	Average 45%
		Grand Average 48%

AFCRL Channel Vocoder - 2400 bits/sec

PB List No.	% Correct	
1E	80	
2E	70	
6E	70	
7E	70	Average 73%

It should not be deduced from these results that the pattern matching technique does not have promise:

a. The system tested is still in its early experimental stages and can undoubtedly be somewhat improved.

b. The channel vocoder with which it was paired provides only marginally satisfactory intelligibility when operated by itself. There was probably a significant saving in information rate required for equal intelligibility. At the present time this could be demonstrated only with the system operating with a greater pattern storage capacity, or with the pattern matching system operating with a vocoder that provides adequate speech intelligibility at a lower bit rate than the vocoder presently used with the AFCRL pattern matching system.

3.1.4 PB word scores -- female talkers

The results obtained with the female talkers were presented in Fig. 4. Inasmuch as none of the systems tested were designed for use with female talkers the results are, perhaps, only of academic interest. The engineers responsible for the various vocoders tested pointed out that the pitch-tracking circuits used in their instruments were designed specifically for male voices but could probably be modified to operate more effectively with female voices.

Nevertheless, in view of the generally lower intelligibility scores for the female voices than the male, even on the reference system, the design of a vocoder that will provide adequate intelligibility for both male and female voices is probably a challenging task.

4. EXPERIMENTATION WITH COMBINED "PEAK-PICKER" AND SPECTRUM SAMPLING (NARROW-BAND) SYSTEM

One of the simplest speech compression techniques studied under this contract was the so-called spectrum sampling (narrow-band) system. This system, being researched and developed under a U. S. Army Contract (DA-36-039-SC-78078) was available for further experimentation at Bolt Beranek and Newman Inc.

Although, as seen in Fig. 2, the particular narrow-band system available for testing did not perform as well as might be hoped, it provided a convenient and ready means for studying how well two speech bandwidth compression techniques might be combined to provide a greater reduction in the bandwidth of processed speech.

One technique which should markedly improve the performance of the narrow-band system would be to make one or more of the filters "track" those spectral peaks that are of sufficient relative amplitude to clearly distinguish them as vocal tract resonances. For vowel sounds, these spectral peaks would be the formants -- hence, the title "formant tracker."

Normally, each of the six crystal filters in the narrow-band system is centered about a fixed location in the lower side band of a suppressed-carrier modulated signal, where the modulating signal is broad-band speech and the carrier is derived from a crystal-controlled oscillator. To move the filter about in the audio spectrum simply requires a change in the carrier frequency. A typical channel is shown in Fig. 5.

Figure 6 is a block diagram of the control system for one of the six narrow-band system filters. The input, a wide-band speech signal, is

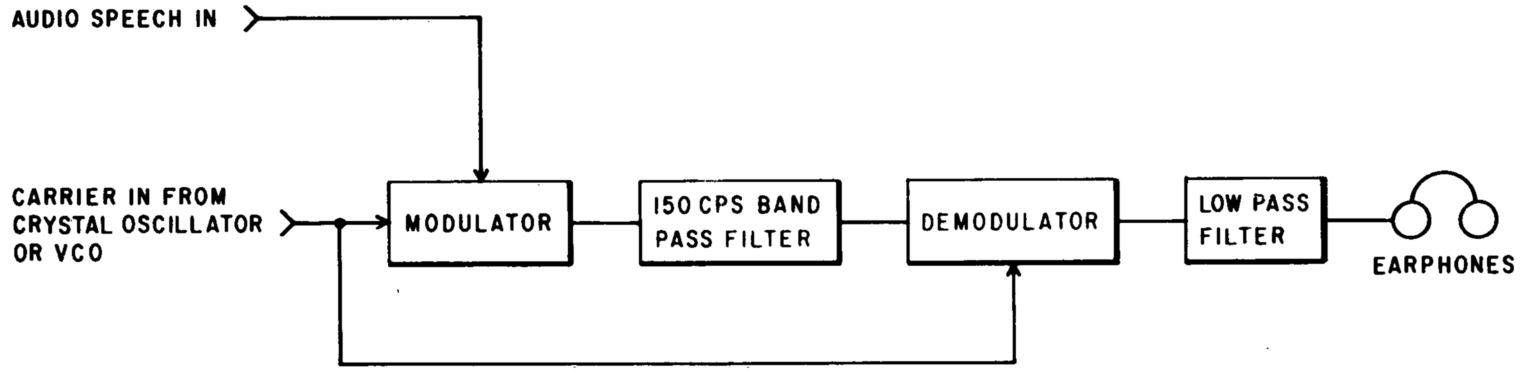

FIG. 5 BLOCK DIAGRAM OF ONE CHANNEL OF NARROW-BAND SPEECH FILTERING SYSTEM

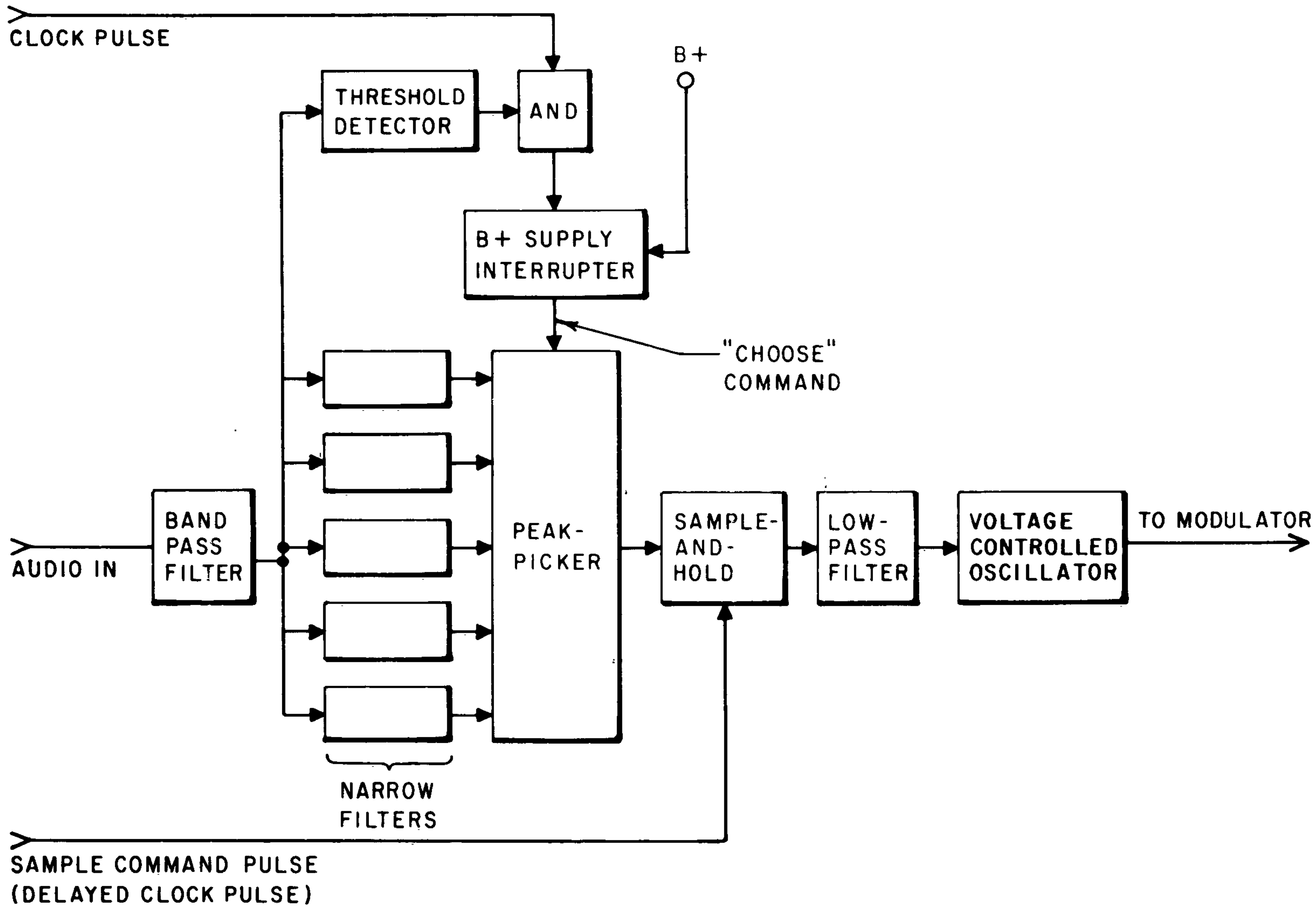

FIG.6 BLOCK DIAGRAM OF FORMANT TRACKER

first bandpass-filtered to establish the limits over which the narrow-band system filter will be made to track. The resulting band-limited signal is fed to a bank of five narrow filters, each of which has an associated detector. The detector outputs are fed to a "peak-picker," whose function it is to determine which of the five narrow filters exhibits the largest output at a given (clock) time.

The peak-picker output is, in the interval between clock times, a steady DC voltage proportional to the number (e.g., 1, 2, 3, 4 or 5) of the narrow filter with the largest output. The peak-picker output is completely independent of the actual voltages presented at its inputs over a range of input voltages in excess of 40 dB.

The peak-picker is forced, whenever the level of the band-limited speech exceeds some threshold, to choose at clock time the largest of its five inputs. Since the clock rate is approximately 400 pps, the peak-picker output is a time-varying voltage which may change every 1/400 second, and which can assume one of five discreet levels. The presence of the lowest output level indicates, in general, the presence of a spectral maximum located more within the passband of the first narrow filter than any of the other four. The next higher output level indicates a (relative) spectral maximum in the band covered by the second narrow filter, etc.

The peak-picker requires a finite time in which to make its decision, and so its output is examined by means of a sample-and-hold circuit which is strobed only after the peak-picker output has assumed a steady value. Since the sampling command pulse disappears before the next clock signal arrives at the peak-picker, the sample-and-hold output contains none of the "return-to-zero" transients present in the peak-picker output.

The sample-and-hold output is fed via a smoothing filter to a voltage-controlled oscillator (VCO), which serves as the modulation system carrier (in place of the crystal oscillator described above).

To avoid random shifting of the VCO frequency (i.e., of the peak-picker output) when there are no appreciable spectral peaks, the "choose" command to the peak-picker (an interruption of the B^+ voltage) is logically "ANDed" with the output of a threshold detector. Since the threshold detector measures the overall level at the output of the band-limiting filter, the peak-picker is not forced to find a spectral maximum (which it might do in a random manner) unless the probability is high that one really exists.

The decay time constant of the sample-and-hold circuit is adjusted so that, in the absence of spectral maxima above threshold, the control voltage presented to the VCO slowly decays from the last definite value to zero.

"Peak-picker" circuits. Figure 7 is the circuit diagram of the formant tracker designed to track, for preliminary exploratory purposes, a narrow-band system crystal filter over the audio range from 1000 to 1800 cps. Trackers for other frequency ranges would be identical except for component values in the bandpass and narrow filters.

The peak-picker, similar to a system by Flanagan,* consists of five gas thyratrons arranged in an "exclusive-OR" configuration. The grid of each tube is supplied with the sum of a positive-going ramp and

* J. L. Flanagan, "Automatic Extraction of Formant Frequencies from Continuous Speech," J. Acoust. Soc. Amer., 1, 110-118, 1956.

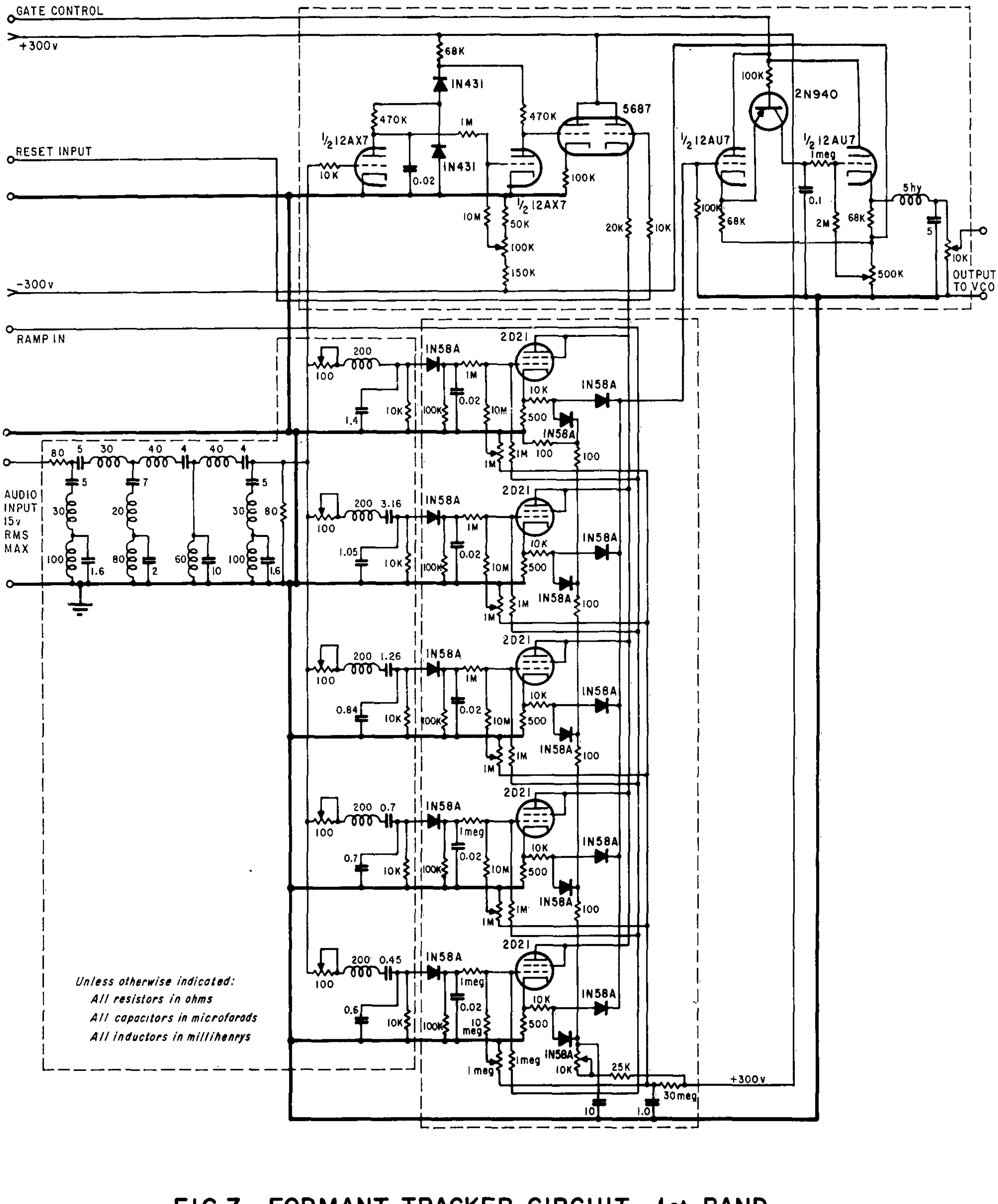

FIG. 7 FORMANT TRACKER CIRCUIT, 1st BAND

the detected output of one of the narrow filters. If all tubes are adjusted for equal firing thresholds, then as the ramp rises the first tube to fire will be the one whose associated narrow filter exhibits the largest output. Since all tubes share a common plate load resistor, the firing of one tube precludes the firing of any of the others until the plate supply voltage is removed. This "choose" command occurs only if the threshold detector indicates the presence of sufficient energy in the bandpass filter to warrant making a determination.

When a tube fires, its plate current is relatively independent of the grid voltage, and so the voltage developed across any cathode load will be a steady DC level. The wiper of the potentiometer in the cathode circuit of the first tube is adjusted for, say, a 1-volt DC output when that tube is fired, the pot of the second tube is set for 2 volts out, etc. All five potentiometer outputs are added by means of a diode adder, and since only one tube fires at a time, its output will be 1, 2, 3, 4 or 5 volts. This quantized signal is fed via a cathode follower to the sample-and-hold circuit described above.

A breadboard model of the peak-picker described above was made to track, over the frequency range 1000 to 1800 cps, the effective center frequency of a filter having, at the 30 dB downpoints, a width of 150 cps. The filter was always presumably located in the frequency region having the greatest concentration of energy.

Results. Because of the narrow bandwidth of the single filter, the speech signal was nearly unintelligible and could not be tested by the usual intelligibility test techniques. However, it was the judgment of the listeners that the output of the filter was more intelligible and more speech-like when the filter tracked the "peak" energy in the range of 1000 to 1800 cps than when the filter was fixed at any location in the same range.

The results were deemed sufficiently encouraging to undertake the breadboarding of additional peak-pickers to control the position of two other narrow-band filters over other frequency regions. The speech from a three-band system should provide speech of sufficient intelligibility to be measurable.

At this time it became apparent that the funds remaining in the contract would not permit both the completion of this combined peak-picking/narrow-band system and the intelligibility testing of other speech compression systems being developed by Air Force contractors during the summer and fall of 1962. The latter tests were judged to be more important and the experimentation started on the peak-picking/narrow-band system was terminated.

The peak-picking/narrow-band system does represent a somewhat different technique for achieving the bandwidth compression of speech than any of the other systems tested. It is our impression, however, from the brief tests we made, that this sytem has but limited promise as a low information rate speech transmission system.

Rome Air Development Center, Griffiss AF Base, New York
Report No. RADC-TDR-63-90, AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES. Tech. Report, 23 February 63, 18 p., 7 figures.
Unclassified Report

The results of previous technical reports, prepared under this contract, are summarized and results of tests of various recently developed speech compression systems are presented and analyzed. The results can be summarized as follows: (a) semi-vocoders, operating at 9600 bits/sec and channel vocoders, at 2400 bits/sec, will provide speech of adequate intelligibility and quality for most military communications. The voice quality of the semi-vocoders will usually be somewhat superior to that of the channel vocoders. Results with as yet incomplete systems suggest that the Tarasoff-Daguet technique and the narrow-band spectrum sampling technique will provide speech intelligibility and quality comparable to that of the semi-vocoder, with about 9600 bits/sec channel capacity during transmission; (b) formant-tracking vocoders operating at about 1200 bits/sec can probably be developed to the point where they will provide speech intelligibility comparable to that from the channel vocoders operating at 2400 bits/sec; (c) transmitting a restricted set of speech "patterns," from a larger set obtained from a digitized vocoder-type of speech analyzer-synthesizer, will provide some reduction in the information transmission rate normally required by that particular vocoder. Although this "pattern" matching technique is still in its early experimental stages it appears that the normal bit rate for a given vocoder system may, by this technique, be reduced to 2/3 and possibly to 1/2 its normal magnitude.

1. Speech transmission
2. Systems for bandwidth compression of speech
3. Channel vocoder
4. Semi-vocoder
5. Formant-tracking vocoder
6. Narrow-band speech system

I. Project No. 4519, Task No. 45350
II. Contract AF30(602)-2235
III. Bolt Beranek and Newman Inc., Cambridge, Mass.
IV. K. D. Kryter and J. H. Ball
V. BBN Report No. 978
VI. TDR-63-90
VII. In ASTIA collection

Rome Air Development Center, Griffiss AF Base, New York
Report No. RADC-TDR-63-90, AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES. Tech. Report, 23 February 63, 18 p., 7 figures.
Unclassified Report

The results of previous technical reports, prepared under this contract, are summarized and results of tests of various recently developed speech compression systems are presented and analyzed. The results can be summarized as follows: (a) semi-vocoders, operating at 9600 bits/sec and channel vocoders, at 2400 bits/sec, will provide speech of adequate intelligibility and quality for most military communications. The voice quality of the semi-vocoders will usually be somewhat superior to that of the channel vocoders. Results with as yet incomplete systems suggest that the Tarasoff-Daguet technique and the narrow-band spectrum sampling technique will provide speech intelligibility and quality comparable to that of the semi-vocoder, with about 9600 bits/sec channel capacity during transmission; (b) formant-tracking vocoders operating at about 1200 bits/sec can probably be developed to the point where they will provide speech intelligibility comparable to that from the channel vocoders operating at 2400 bits/sec; (c) transmitting a restricted set of speech "patterns," from a larger set obtained from a digitized vocoder-type of speech analyzer-synthesizer, will provide some reduction in the information transmission rate normally required by that particular vocoder. Although this "pattern" matching technique is still in its early experimental stages it appears that the normal bit rate for a given vocoder system may, by this technique, be reduced to 2/3 and possibly to 1/2 its normal magnitude.

1. Speech transmission
2. Systems for bandwidth compression of speech
3. Channel vocoder
4. Semi-vocoder
5. Formant-tracking vocoder
6. Narrow-band speech system

I. Project No. 4519, Task No. 45350
II. Contract AF30(602)-2235
III. Bolt Beranek and Newman Inc., Cambridge, Mass.
IV. K. D. Kryter and J. H. Ball
V. BBN Report No. 978
VI. TDR-63-90
VII. In ASTIA collection

Rome Air Development Center, Griffiss AF Base, New York
Report No. RADC-TDR-63-90, AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES. Tech. Report, 23 February 63, 18 p., 7 figures.
Unclassified Report

The results of previous technical reports, prepared under this contract, are summarized and results of tests of various recently developed speech compression systems are presented and analyzed. The results can be summarized as follows: (a) semi-vocoders, operating at 9600 bits/sec and channel vocoders, at 2400 bits/sec, will provide speech of adequate intelligibility and quality for most military communications. The voice quality of the semi-vocoders will usually be somewhat superior to that of the channel vocoders. Results with as yet incomplete systems suggest that the Tarasoff-Daguet technique and the narrow-band spectrum sampling technique will provide speech intelligibility and quality comparable to that of the semi-vocoder, with about 9600 bits/sec channel capacity during transmission; (b) formant-tracking vocoders operating at about 1200 bits/sec can probably be developed to the point where they will provide speech intelligibility comparable to that from the channel vocoders operating at 2400 bits/sec; (c) transmitting a restricted set of speech "patterns," from a larger set obtained from a digitized vocoder-type of speech analyzer-synthesizer, will provide some reduction in the information transmission rate normally required by that particular vocoder. Although this "pattern" matching technique is still in its early experimental stages it appears that the normal bit rate for a given vocoder system may, by this technique, be reduced to 2/3 and possibly to 1/2 its normal magnitude.

1. Speech transmission
2. Systems for bandwidth compression of speech
3. Channel vocoder
4. Semi-vocoder
5. Formant-tracking vocoder
6. Narrow-band speech system

I. Project No. 4519, Task No. 45350
II. Contract AF30(602)-2235
III. Bolt Beranek and Newman Inc., Cambridge, Mass.
IV. K. D. Kryter and J. H. Ball
V. BBN Report No. 978
VI. TDR-63-90
VII. In ASTIA collection

Rome Air Development Center, Griffiss AF Base, New York
Report No. RADC-TDR-63-90, AN EVALUATION OF SPEECH COMPRESSION TECHNIQUES. Tech. Report, 23 February 63, 18 p., 7 figures.
Unclassified Report

The results of previous technical reports, prepared under this contract, are summarized and results of tests of various recently developed speech compression systems are presented and analyzed. The results can be summarized as follows: (a) semi-vocoders, operating at 9600 bits/sec and channel vocoders, at 2400 bits/sec, will provide speech of adequate intelligibility and quality for most military communications. The voice quality of the semi-vocoders will usually be somewhat superior to that of the channel vocoders. Results with as yet incomplete systems suggest that the Tarasoff-Daguet technique and the narrow-band spectrum sampling technique will provide speech intelligibility and quality comparable to that of the semi-vocoder, with about 9600 bits/sec channel capacity during transmission; (b) formant-tracking vocoders operating at about 1200 bits/sec can probably be developed to the point where they will provide speech intelligibility comparable to that from the channel vocoders operating at 2400 bits/sec; (c) transmitting a restricted set of speech "patterns," from a larger set obtained from a digitized vocoder-type of speech analyzer-synthesizer, will provide some reduction in the information transmission rate normally required by that particular vocoder. Although this "pattern" matching technique is still in its early experimental stages it appears that the normal bit rate for a given vocoder system may, by this technique, be reduced to 2/3 and possibly to 1/2 its normal magnitude.

1. Speech transmission
2. Systems for bandwidth compression of speech
3. Channel vocoder
4. Semi-vocoder
5. Formant-tracking vocoder
6. Narrow-band speech system

I. Project No. 4519, Task No. 45350
II. Contract AF30(602)-2235
III. Bolt Beranek and Newman Inc., Cambridge, Mass.
IV. K. D. Kryter and J. H. Ball
V. BBN Report No. 978
VI. TDR-63-90
VII. In ASTIA collection