AD-A095 664 LERNER TECHNOLOGY INC WILLIAMSVILLE NY F/G 5/8
AUTOMATED INTELLIGIBILITY MEASUREMENT SYSTEM.(U)
NOV 80 T LERNER F30602-80-C-0065
UNCLASSIFIED RADC-TR-80-353 NL

1 OF 1
AD
A095664

END
DATE
FILMED
3-81
DTIC

(12)

SC

RADC-TR-80-353
Final Technical Report
November 1980

AD A095664

AUTOMATED INTELLIGIBILITY MEASUREMENT SYSTEM

LEVEL II

Lerner Technology, Inc.

Theodore Lerner

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

This effort was funded by the Laboratory Director's Fund.

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, New York 13441

81 2 27 059

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-80-353 has been reviewed and is approved for publication.

APPROVED: Donald M Ottinger

DONALD M. OTTINGER, Capt, USAF
Project Engineer

APPROVED: Fred I Diamond

FRED I. DIAMOND, Technical Director
Communications and Control Division

FOR THE COMMANDER: John P. Huss

JOHN P. HUSS
Acting Chief, Plans Office

This effort was funded by the Laboratory Director's Fund.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-80-353	2. GOVT ACCESSION NO. AD-A095 664	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AUTOMATED INTELLIGIBILITY MEASUREMENT SYSTEM		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report 14 Dec 79 — 13 Aug 80
		6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) Theodore Lerner		8. CONTRACT OR GRANT NUMBER(s) F30602-80-C-0065
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lerner Technology, Inc. 195 N. Long Street Williamsville NY 14221		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61101F LD0110C1
11. CONTROLLING OFFICE NAME AND ADDRESS Rome Air Development Center (DCID) Griffiss AFB NY 13441		12. REPORT DATE November 1980
		13. NUMBER OF PAGES 43
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Same		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: Capt Donald M. Ottinger (DCID)

This effort was funded by the Laboratory Directors Fund.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Speech
Speech Intelligibility

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report documents the technical effort conducted under Contract F30602-80-C-0065 and funded by the Laboratory Director's Fund. The report discusses the design, development, and test of a microprocessor based system for measuring the intelligibility of speech processed through a communications channel. The major components of the intelligibility measurement system include an audio test tape of fifty phonetically balanced (PB) words, a bank of twelve audio bandpass

(Cont'd)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

394

Item 20 (Cont'd)

cont

filters, an audio power detector, and a microprocessor. A time-frequency matrix of the starting and ending phonemes of each PB word is stored in memory as the reference. The system works by inputting a recording of the audio test tape as received over a communications channel. Timing marks included within the audio test tape allow comparison of the time-frequency history of the test tape. This comparison of the reference values with the test data yields an intelligibility score. Initial tests of the automatic intelligibility measurement system promise performance which will be superior to that of existing automated techniques.

CONTENTS

List of Figures

List of Tables

EVALUATION

The objective of this effort was to investigate a technique for automatically measuring the intelligibility of speech processed over a communications channel. A breadboard model was built and successfully demonstrated the ability to measure speech intelligibility.

Several efforts have attempted to produce an automated speech intelligibility measurement device, but very few have produced effective hardware. The technique evaluated here promises to overcome the failures of many past efforts. The results presented here should help lead to a final solution of the problem of efficiently and objectively evaluating the performance of voice communications systems.

Donald M. Ottinger

DONALD M. OTTINGER
Project Engineer

1. INTRODUCTION

There are a number of criteria for evaluating voice modems. These include Intelligibility - which refers to the ability to understand words, Quality - which refers basically to system acceptance, Listener Fatigue, and Speaker Recognition. Of these, the most important is Intelligibility. Many tests have been developed for measuring Intelligibility. These involve transmitting words, syllables, phrases or sentences through the system, and having a panel of listeners try to identify what was transmitted. The results are scored and an intelligibility score is then determined. Tw tests that have found wide acceptance are the Modified Rhyme Test and the PB-50 Test. (An excellent discussion of speech intelligibility and quality testing is given in Section 8 of Reference (1).)

Intelligibility tests are expensive, often difficult to reproduce, and sometimes inconclusive. To get around these problems a number of analytic and semi-analytic techniques have been developed. An example of an analytic approach to the problem is given in Reference (2). A semi-analytic technique that is often used is based on the Articulation Index. This is instrumented in such equipment as VIAS or SCIM by transmitting a "speech-like" test signal through the system and measuring test signal to noise at the output in a number of bands. The Articulation Index is then computed automatically by the equipment. Where test tone to noise measurements have been made on the equipment, Articulation Index can be computed through the use of weighting of test tone to

noise as a function of the particular band. The method for computing the articulation index is given in Reference (3). Intelligibility scores can then be estimated from the articulation index.

While these analytic and semi-analytic techniques are generally less expensive and more repeatable than the intelligibility test, they often give incorrect results. The analytic and semi-analytic techniques are based on certain assumptions of the nature of the modem. These assumptions are often not valid and there is considerable discrepancy between articulation index scores and intelligibility scores for different types of modems.

The techniques based on test tone to noise measurements or articulation index were developed primarily for use on essentially linear system such as amplitude modulation or standard frequency modulation. Many modern systems are basically non-linear in nature. For example, in delta modulation, slope overload will cause a high frequency signal component to be lost in the presence of a relatively strong low frequency component. A low level signal component will also be masked by the quantization noise. On the other hand, a median level high frequency signal will be passed well by the delta modulation system. Various speech phonemes fall into each of these categories, namely, some are median level high frequency signals, some are low level signals, and some are made up of relatively strong low frequency signals with high frequency components. The techniques based on test tone to noise measurements or articulation index do not take these facts into account and therefore give erroneous results.

Another example of a non-linear technique that is used in many modern systems to improve intelligibility at low C/N_o is speech signal compression. Signal compression enhances low level phonemes, but introduces high frequency components that may not have been present in the original signal. A technique that does not take both of these facts into account will not yield a correct score.

Certain phonemes, in particular, the stop consonants such as "t" or "p" depend for recognition on the time history of the sound at least as much as they do on the frequency composition. A number of voice communication systems do not perform well for these types of phonemes. These include some of the vocoders when operated at low data rate, and some of the TASI type systems. The intelligibility measuring techniques based on articulation index do not take time history into account, and thus give erroneous results for these voice communication systems.

A system is described in this report which, by using actual speech phonemes, overcomes many of the problems that occur in existing intelligibility measuring instruments. In particular the system is based on the use of a test tape on which is recorded a set of PB words. The starting and ending phonemes of each word are located in time by means of precise timing signals. Prior to the use of the tape for the evaluation of speech modems, the tape is played through a signal analysis unit which measures for each of the phonemes, the energy content in a set of selected frequency levels and during a set of selected time periods. Thus a time-

frequency matrix is obtained which is characteristic of each phoneme. This time frequency matrix which serves as a reference is stored in a microcomputer.

The tape can then be played thru a modem system for which an intelligibility score is desired. The output of the system is then recorded on tape. The resulting tape is then played through the signal analyzer unit which obtains a new time frequency matrix for each starting and ending phoneme. In obtaining the time frequency matrix the precise timing signal is used to insure that the new time frequency matrix is obtained at exactly the same time as the reference time frequency matrix. The new time frequency matrix, which is subject to the noise and distortion of the modem, is fed to the microcomputer, where it is compared with the reference, and an intelligibility measure for each phoneme is computed. The intelligibility score is then computed by averaging with suitable weighting over all of the phonemes.

The instrumentation for voice intelligibility measurement described in this report promises performance which will be superior to that of existing techniques. The most important reason for the improved performance is the fact that actual speech phonemes are used, so that the effect of noise or distortion is more realistically taken into account. Since actual speech sounds are used, the signal statistics are realistic, and the instrumentation is usable for measuring performance of modern modem systems which are based on speech signal statistics. Further, the system is sufficiently flexible so as to permit modifications to be made in the

intelligibility estimation algorithms as new knowledge is obtained concerning speech intelligibility.

2. SYSTEM DESCRIPTION

2.1 Overall Description. The complete system consists of two major components: the test tape and the scoring system. The test tape contains a set of 50 phonetically balanced words each of which is preceded by a timing code. A block diagram of the scoring system is shown in Figure 1.

Prior to delivery, information concerning the frequency content of the starting and ending phoneme of each of the 50 words is stored in the computer memory for reference. This is accomplished in the following manner. The test tape is played through the system. Under computer control the code preceding each word is decoded, and the computer then times to the first sample measurement. This time is, in general, different for each word, and has been previously established by analysis. At the sample time the content of each of the filters shown in Figure 1 is sampled successively by the multiplexer. Each of the 12 filter samples is converted to an 8 bit digital word and is stored in computer memory. The computer then times to the next sample and the filter sampling process is repeated. The number of sample measurements per phoneme is, in general, different for each phoneme. This number has been determined by analysis and is stored in computer memory. The entire process is repeated for each word starting in each case with the decoding of the timing code. At the end of this calibration, the computer memory contains the following information: the number of samples in each phoneme for which measurements are to be made; the time from the timing code to each sample; and the frequency content of each sample.

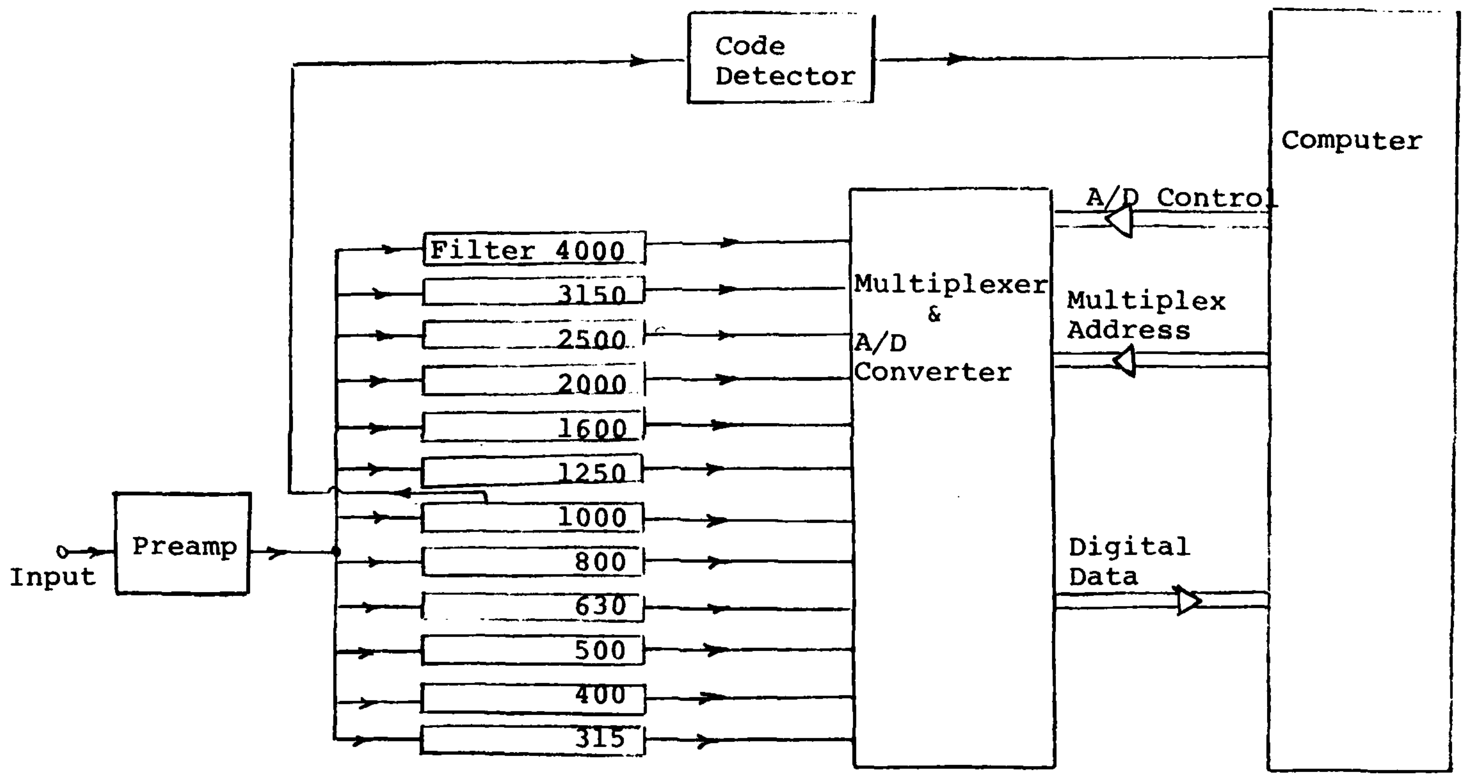

Figure 1
Block Diagram - Intelligibility Scoring System

To evaluate a communications system, the test tape is played through the communication system and a new tape made from the output. The new tape contains the timing codes and words of the original test tape with whatever distortion and noise has been added by the communication system.

The new tape is then played through the Intelligibility Scoring System. The operation is now similar to that for calibration. The timing code is decoded, the system times to each sample time, and at each sample time it measures and stores in memory the contents of each of the 12 filters. Because of the noise and distortion introduced by the communication system, the filter content will, in general, be different from that obtained with the original test tape.

After each of the 50 words has been analyzed as described above, an intelligibility measure is obtained for each phoneme, an intelligibility score is obtained for each word, and the word intelligibility scores are then averaged to obtain an overall intelligibility score for the communication system.

The intelligibility measure is a simple distance measure, and is computed as follows:

$$\text{Intelligibility Measure} = \frac{(a_1 b_1 + a_2 b_2 + \ldots\ldots)^2}{(a_1^2 + a_2^2 + \ldots\ldots)(b_1^2 + b_2^2 + \ldots\ldots)}$$

where: a_i is the reference content of the i^{th} filter obtained during the calibration.

b_i is the content i^{th} filter obtained during the test run.

The summation is over the 12 filters if only a single sample is measured for the phoneme. If more than one sample is measured for the phoneme, the summation will be over 24 or 36 depending on whether two or three samples are measured for the phoneme.

The intelligibility score for each phoneme is obtained from a linear relationship between intelligibility score and intelligibility measure. The parameters for this linear relationship were obtained by comparing the test scores obtained with human subjects for each phoneme with the corresponding intelligibility measure for that phoneme. In general, the relationship is different for each phoneme. An intelligibility score for the word is obtained by selecting the lower of the two intelligibility scores obtained for the starting and ending phonemes.

The overall intelligibility score is obtained by simply averaging over the word intelligibility scores.

2.2 Test Tape. As mentioned previously, the test tape contains a set of 50 phonetically balanced words each of which is preceded by a timing code. The timing code consists of a 1 KHz tone which is modulated by a pulse pattern as shown in Figure 2.

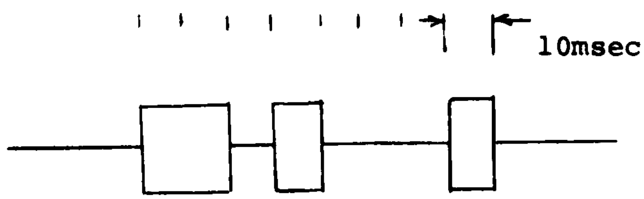

Figure 2. Timing Code Pulse Pattern

This code has a number of properties which make it useful for this application. It can provide accurate timing because it has good autocorrelation properties. The autocorrelation function is 4 for zero shift and is no higher than 1 for any pulse shift other than zero. The code has sufficient energy to insure good decoding under signal noise ratios lower than those for which usable speech communication can be obtained. The frequency components of the code being centered on 1 KHz are such that they will pass any reasonable speech communication system.

Tests conducted on the code with decoder circuitry and algorithm described elsewhere in the report yielded the following results:

1. Reliable decoding was obtained with code signal amplitudes between 2 volts and 12 volts.
2. Reliable decoding in noise was obtained at a value of code signal/noise spectral density of 37 dbHz. Approximately 50% decoding was obtained at a value of code signal/noise spectral density of 31 dbHz. (This is well below the value at which usable speech intelligibility is obtained.)
3. Total jitter in decode delay was $\pm$2.5 msec under all conditions.

2.3 System Hardware. In this section a detailed description is presented of each of the major blocks shown in the block diagram of Figure 1.

2.3.1 Pre-Amplifier. A schematic of the pre-amplifier is shown in Figure 3. The purpose of the pre-amplifier is to amplify the signal from the tape recorder to a level suitable for driving the filters. As shown, the preamplifier consists of two stages with an overall gain from the audio input of 30. A volume control is provided to permit adjusting the signal level out of the pre-amplifier to the correct value of approximately 6 volts peak. It should be noted that, because of the normalizing properties of the intelligibility measure algorithm, the adjustment is not critical. In addition to the audio input which is normally used, a second input with switchable gain is provided to permit adding noise for system test and evaluation.

2.3.2 Filters. As shown in Figure 1, twelve filters are provided, each with a different center frequency. Each filter has a bandwidth approximately 20% of the center frequency. A schematic of the filter circuitry is shown in Figure 4. The filter proper is made up of the first two stages which are stagger tuned, the first stage 5% below the center frequency, the second stage 5% above the center frequency. The tuning of the filters is accomplished by selection of the resistors, R_{a1}, R_{a2}, R_{a3}, R_{b1}, R_{b2}, R_{b3} and the capacitors, C. The The values of these resistors and capacitors for each filter are given in Table 1. Stagger tuning provides a relatively flat response in the passband of the filter with good rejection outside the passband. A sample response for the 1 KHz filter is shown in Figure 5.

The third and fourth stage make up a full-wave rectifier and single pole low pass filter. The fifth

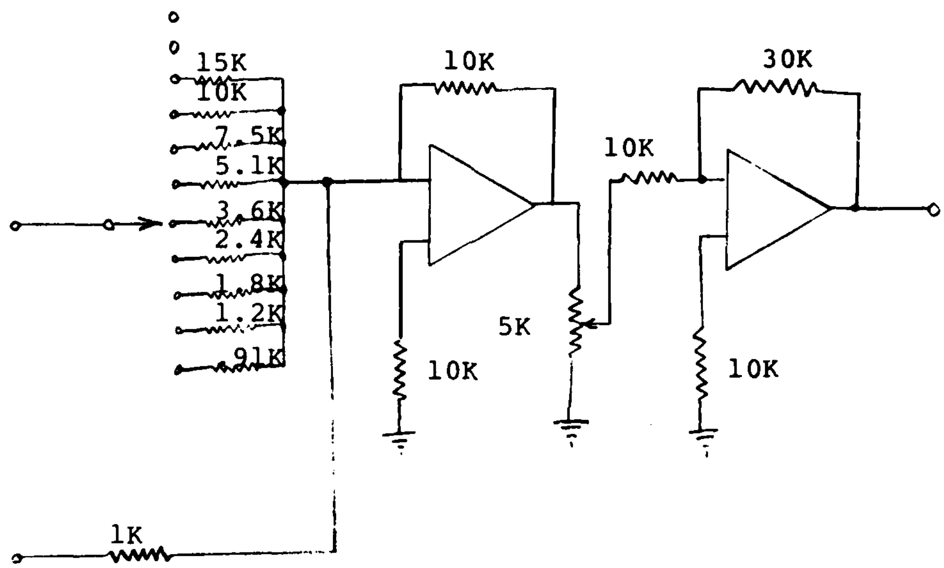

Figure 3. Preamplifier

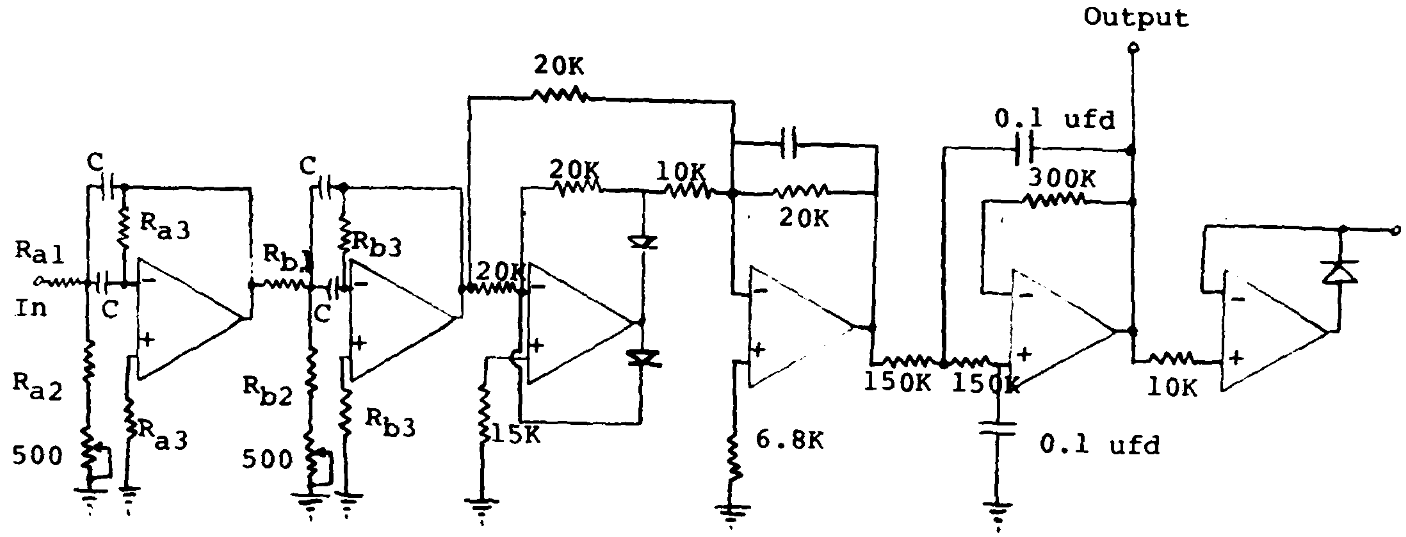

Figure 4. Filter Schematic

Table 1. Selected Values for Filter Components

f_o	f_{oa}	f_{ob}	C	R_1	R_2	R_3
250	237.5		.047	68K	470	300K
		262.5	.047	62K	390	270K
315	299.25		.047	56K	330	220K
		330.75	.047	51K	270	200K
400	380		.047	43K	220	180K
		420	.047	39K	150	160K
500	475		.047	36K	100	150K
		525	.047	33K	82	130K
630	598.5		.01	130K	1000	510K
		661.5	.01	120K	1000	470K
800	760		.01	100K	820	430K
		840	.01	91K	680	390K
1000	950		.01	82K	560	330K
		1050	.01	75K	470	300K
1250	1187.5		.01	68K	390	270K
		1312.5	.01	62K	330	240K
1600	1520		.01	51K	270	200K
		1680	.01	47K	220	180K
2000	1900		.0047	91K	680	360K
		2100	.0047	82K	560	330K
2500	2375		.0047	68K	470	300K
		2625	.0047	62K	390	270K
3150	2992.5		.0047	56K	330	220K
		3307.5	.0047	51K	270	200K
4000	3800		.0047	43K	220	180K
		4200	.0047	39K	150	160K

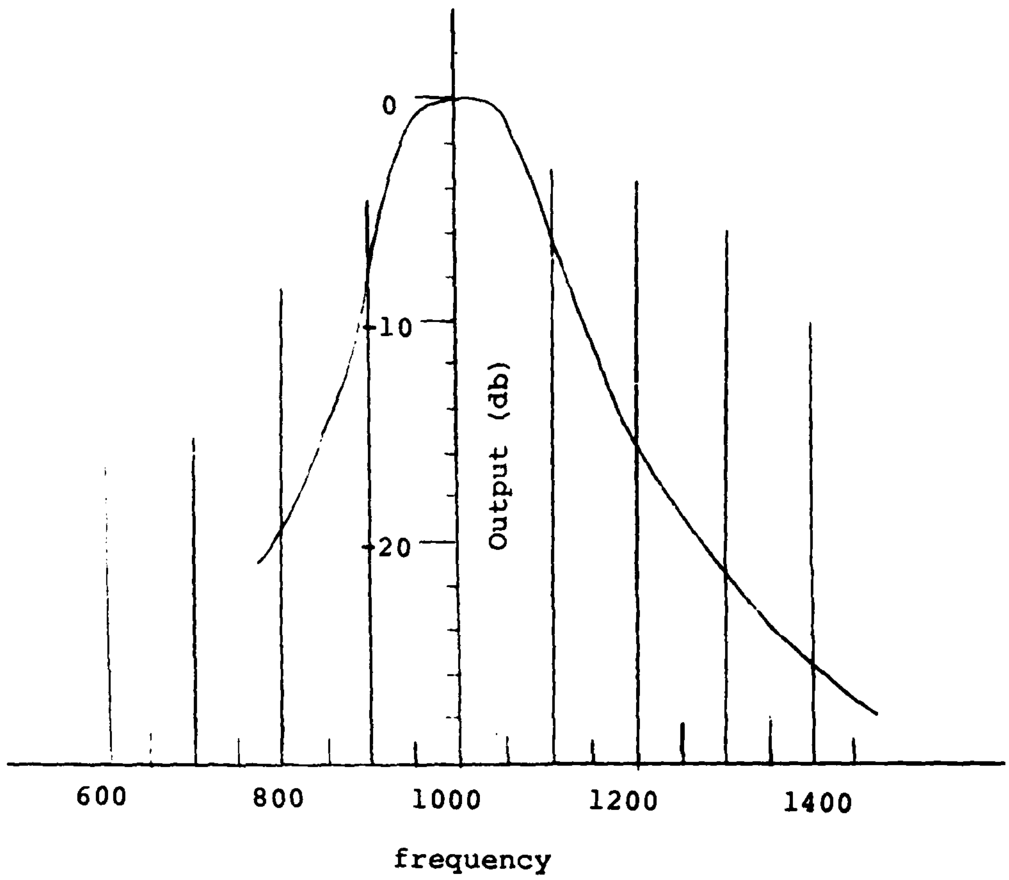

Figure 5. Frequency Response of 1KHz Filter

stage is used for a two pole low pass filter. Thus the overall low pass response is that of a three pole filter. Each pole is tuned to approximately 11 Hz. The use of this three pole filter provides good smoothing with essentially independent measurements when the measurements are separated by 50 milliseconds or more.

2.3.3 Code Detector. The code detector works with the 1 KHz filter (as shown in Figure 1) to provide a demodulated signal to the computer for accurate decoding. A schematic of the code detector is shown in Figure 6. The signal to the code detector is taken from the second stage of the 1KHz filter. As shown the first two stages of the code detector constitute a full wave rectifier and low pass filter. The response time of this low pass filter is faster than that of the 1 KHz filter in order to insure accurate timing. The final stage of the code detector is a comparator which provides a signal at 0 volts in the absence of a code pulse and +5 volts in the presence of a code pulse. These signals are suitable for driving the computer for decoding.

2.3.4 Multiplexer and A/D Converter. The multiplexer and A/D converter is a commercial unit, the AIM16, built by Connecticut Microcomputer, Inc. The AIM16 is capable of selecting 1 of 16 analog inputs in response to a digital multiplex address. The selected analog input signal is converted to an 8 bit digital signal which is available to the computer. The conversion time is less than 100 microseconds. The input voltage range is 0 to 5.12 volts which is converted to a count between

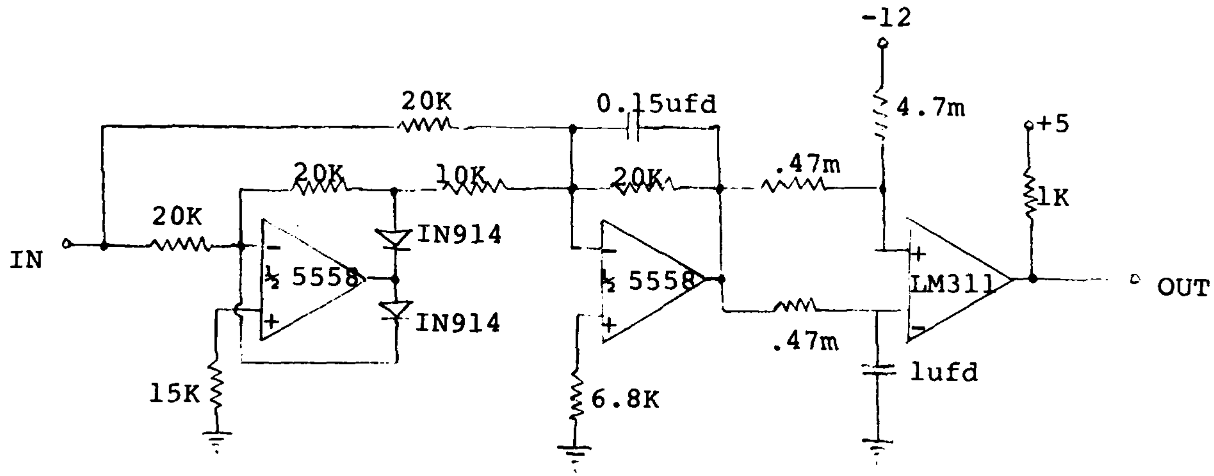

Figure 6. Code Detector

o and 255 (00 to FF hex). Resolution is thus 20 millivolts per count.

2.3.5 Computer. The computer used is the AIM65 built by Rockwell International Corporation. It is based on the 6502 microprocessor. In addition to the 4K of RAM memory resident on the computer board, an additional 8K of RAM memory was added to permit adequate storage of programs and data. The computer permits very simple interface to external circuitry through a set of four 8 bit I/O ports, and thus was particularly useful for this application where such ease of interfacing was an important consideration. The computer can be programmed in machine language and in Basic through the use of a Basic RAM.

2.4 System Software. The system software is made up of a number programs written both in Basic and machine language. The master program, written in Basic, is SCOR2. This in turn calls as a subroutine STOR8 which is written in machine language. STOR8 calls as subroutines DCOD5 and SAMP3. Listings for these programs are given in Tables 2, 3, 4 and 5.

DCOD5 performs the decoding of the timing code. It samples the output of the code detector every 50 milliseconds. The successive outputs are shifted into memory so that 19 successive time samples are stored. These memory locations are then examined for the presence of a proper code by adding the numbers stored in locations corresponding to the presence of code pulses and subtracting the numbers corresponding to the absence of code pulses. If this process yields 3 or more a code is

declared to be present. This permits 1 pulse of the code to be lost due to noise, or it permits one false pulse to occur.

SAMP3 addresses the multiplexer and A/D converter so as to read successively the contents of the 12 filters. These are stored in successive memory locations.

STOR8 performs the initialization of various memory locations as required, it calls DCOD5, and after a timing code is detected, it times to the first sample and then times to succeeding samples of a word. At each sample time it calls SAMP3 to read the contents of the 12 filters. The timing between samples and the number of samples per phoneme have been determined for each word by analysis of the frequency content as a function of time for each word. A listing of the number of samples used in the leading and ending phoneme of each word is given in Table 6. A listing of the times to the samples is given in Table 7. The numbers in Table 7 are multiples of 25 milliseconds and are given in hex. The first number is the time from the code to the first sample, the second number is the time to the second sample, the third number is the time to the third sample, the fourth number is the time to the first sample of the ending phoneme, the fifth number is the time to the second sample of the ending phoneme, and the sixth number is the time to the third sample of the ending phoneme. The hex number 00 indicates that there is no sample. The numbers in Table 7 are in groups of 6 for each word.

SCOR2 calls STOR8 as a subroutine and then performs the necessary calculations to determine an intelligibility score. It is useful to review the function of some of the

lines of SCOR2 to describe the program operation. Lines 20 and 30 identify and call STOR8. Lines 160 through 290 compute the intelligibility measure (A) for each phoneme. Line 120 looks up in memory the number of samples to be used in that computation. Line 330 computes the intelligibility score (TS(M)) for each phoneme. Based on the intelligibility measure (A) and the parameters AH and AL. AH and AL are, in general, different for each phoneme and are found in lines 310 and 320. A listing of AH and AL is given Table 8. Lines 332 through 339 insure that the intelligibility score can never exceed 1 or be less than 0. Lines 360 through 390 select the lowest intelligibility score for each word. Line 410 computes a running average of the intelligibility score.

Table 2. DCOD5

```
3C00  A2  LDX  #13
3C02  A9  LDA  #00
3C04  9D  STA  3C60,X
3C07  CA  DEX
3C08  10  BPL  3C02
3C0A  A9  LDA  #7C
3C0C  8D  STA  A008
3C0F  A9  LDA  #13
3C11  8D  STA  A009
3C14  A2  LDX  #12
3C16  BD  LDA  3C60,X
3C19  9D  STA  3C61,X
3C1C  CA  DEX
3C1D  10  BPL  3C16
3C1F  AD  LDA  A000
3C22  29  AND  #40
3C24  4A  LSR  .A
3C25  4A  LSR  .A
3C26  4A  LSR  .A
3C27  4A  LSR  .A
3C28  4A  LSR  .A
3C29  4A  LSR  .A
3C2A  8D  STA  3C60
3C2D  18  CLC
3C2E  A9  LDA  #00
3C30  6D  ADC  3C63
3C33  6D  ADC  3C6B
3C36  6D  ADC  3C6F
3C39  6D  ADC  3C71
3C3C  38  SEC
3C3D  ED  SBC  3C61
3C40  ED  SBC  3C65
3C43  ED  SBC  3C67
3C46  ED  SBC  3C69
3C49  ED  SBC  3C6D
3C4C  ED  SBC  3C73
3C4F  E9  SBC  #03
3C51  10  BPL  3C5C
3C53  A9  LDA  #20
3C55  2C  BIT  A00D
3C58  F0  BEQ  3C55
3C5A  D0  BNE  3C0A
3C5C  60  RTS
```

Table 3. SAMP3

```
3C80  A9  LDA  #0C
3C82  A8  TAY
3C83  09  ORA  #10
3C85  8D  STA  3CB0
3C88  88  DEY
3C89  A9  LDA  #20
3C8B  8D  STA  A000
3C8E  AD  LDA  3CB0
3C91  8D  STA  A000
3C94  A2  LDX  #49
3C96  CA  DEX
3C97  10  BPL  3C96
3C99  A9  LDA  #20
3C9B  4D  EOR  A000
3C9E  8D  STA  A000
3CA1  AD  LDA  A00F
3CA4  99  STA  27BC,Y
3CA7  CE  DEC  3CB0
3CAA  88  DEY
3CAB  10  BPL  3C89
3CAD  60  RTS
```

Table 4. STOR8

```
3D00  A9  LDA  #00
3D02  8D  STA  A003
3D05  8D  STA  A00B
3D08  8D  STA  3DF8
3D0B  8D  STA  3DF1
3D0E  A9  LDA  #BF
3D10  8D  STA  A002
3D13  A9  LDA  #00
3D15  8D  STA  3CA5
3D18  A9  LDA  #30
3D1A  8D  STA  3CA6
3D1D  A9  LDA  #00
3D1F  8D  STA  3D47
3D22  8D  STA  3D9E
3D25  A9  LDA  #2C
3D27  8D  STA  3D48
3D2A  8D  STA  3D9F
3D2D  20  JSR  3D40
3D30  60  RTS
3D31  4C  JMP  3D00

3D40  20  JSR  3C00
3D43  AE  LDX  3DF8
3D46  BD  LDA  2D26,X
3D49  8D  STA  3DF2
3D4C  A9  LDA  #00
3D4E  8D  STA  3DF3
3D51  A9  LDA  #80
3D53  8D  STA  A008
3D56  A9  LDA  #13
3D58  8D  STA  A009
3D5B  A9  LDA  #20
3D5D  2C  BIT  A00D
3D60  F0  BEQ  3D5D
3D62  EE  INC  3DF3
3D65  AD  LDA  3DF3
3D68  CD  CMP  3DF2
3D6B  D0  BNE  3D51
3D6D  AD  LDA  3DF8
3D70  D0  BNE  3D8D
3D72  A9  LDA  #80
3D74  8D  STA  A000
3D77  A9  LDA  #80
3D79  8D  STA  A008
3D7C  A9  LDA  #13
3D7E  8D  STA  A009
3D81  A9  LDA  #20
3D83  2C  BIT  A00D
3D86  F0  BEQ  3D83
3D88  A9  LDA  #00
3D8A  8D  STA  A000
3D8D  20  JSR  3C80
3D90  EE  INC  3DF8
3D93  AD  LDA  3DF8
3D96  09  CMP  #06
3D98  F0  BEQ  3DB6
3D9A  AE  LDX  3DF8
3D9D  BD  LDA  2D26,X
3DA0  F0  BEQ  3D90
3DA2  18  CLC
3DA3  AD  LDA  3CA5
3DA6  69  ADC  #0C
3DA8  8D  STA  3CA5
3DAB  AD  LDA  3CA6
3DAE  69  ADC  #00
3DB0  8D  STA  3CA6
3DB3  4C  JMP  3D43
3DB6  EE  INC  3DF1
3DB9  AD  LDA  3DF1
3DBC  C9  CMP  #32
3DBE  F0  BEQ  3DF0
3DC0  A9  LDA  #00
3DC2  BD  STA  3DF8
3DC5  18  CLC
3DC6  AD  LDA  3CA5
3DC9  69  ADC  #0C
3DCB  8D  STA  3CA5
3DCE  AD  LDA  3CA6
3DD1  69  ADC  #00
3DD3  8D  STA  3CA6
3DD6  18  CLC
3DD7  AD  LDA  3D47
3DDA  69  ADC  #06
3DDC  8D  STA  3D47
3DDF  8D  STA  3D9E
3DE2  AD  LDA  3D48
3DE5  69  ADC  #00
3DE7  8D  STA  3D48
3DEA  8D  STA  3D9F
3DED  4C  JMP  3D40
3DF0  60  RTE
```

Table 5. SCOR2

```
10  DIM  TS(2)
20  POKE  04, 00: POKE 05, 61
30  Y=USR(0)
40  IN=0
50  NS=12032
60  XS=8192
70  YS=12288
80  AS=11776
90  FOR L=0TO49
100 PRINT! "WORD #"L + 1
110 FOR M=0TO1
120 N=PEEK (NS+2*L+M)
130 S=0
140 SX=0
150 SY=0
160 FOR K=1TON
170 FOR J=0TO2
180 FOR I=0TO3
190 X=PEEK (XS+I)
200 Y=PEEK (YS+I)
210 S=S+X*Y
220 SX=SX+X*X
230 SY=SY+Y*Y
240 NEXT I
250 XS=XS+4
260 YS=YS+4
270 NEXT J
280 NEXT K
290 A=(S*S/SX)/SY
300 PRINT! INT (A*100)/100
310 AH=PEEK (AS+4*L+2*M)
320 AL=PEEK (AS+4*L+2*M+1)
330 TS (M)=10*A/(AH-AL)-AL/(AH-AL)
332 IF TS(M)>1 THEN 337
334 IF TS(M)>0 THEN 339
336 GOTO 340
337 TS(M)=1
338 GOTO 340
339 TS(M)=0
349 PRINT!INT (TS(M)*100)/100
350 NEXT M
360 IF TS(0)>TS(1) THEN 390
370 TS(2)=TS(0)
380 GOTO 410
390 TS(2) = TS(1)
410 IN=(IN*L+TS(2))/
                (L+1)
420 PRINT!INT(IN*100)/
               100
430 NEXT L
```

Table 6. Number of Samples per Phoneme

2F00	02	02	02	02
2F04	02	02	02	02
2F08	02	02	02	02
2F0C	02	02	02	02
2F10	02	02	02	02
2F14	02	02	02	03
2F18	01	02	01	02
2F1C	02	02	02	02
2F20	02	01	01	02
2F24	01	01	02	02
2F28	02	02	02	02
2F2C	01	01	01	02
2F30	01	02	01	01
2F34	02	02	02	01
2F38	02	01	02	02
2F3C	02	01	02	01
2F40	02	01	02	02
2F44	01	01	01	02
2F48	02	01	01	01
2F4C	02	01	01	01
2F50	01	02	01	02
2F54	01	02	01	01
2F58	01	02	01	02
2F50	01	02	01	02
2F60	02	02	01	03

Table 7. Times to Sample Measurements

2C00	31	13	00	40	2C98	00	5E	00	00
2C04	13	00	4E	09	2C9C	96	13	00	59
2C08	00	59	13	00	2CA0	27	00	78	13
2C0C	75	09	00	54	2CA4	00	4A	00	00
2C10	09	00	B4	09	2CA8	A9	1D	00	4F
2C14	00	40	13	00	2CAC	00	00	83	13
2C18	76	09	00	81	2CB0	00	59	13	00
2C1C	09	00	BD	09	2CB4	90	13	00	40
2C20	00	4F	13	00	2CB8	00	00	9E	13
2C24	A2	09	00	68	2CBC	00	4A	00	00
2C28	09	00	BB	09	2CC0	B7	09	00	4A
2C2C	00	4F	13	00	2CC4	09	00	32	13
2C30	A0	09	00	59	2CC8	00	45	13	00
2C34	2C	00	B4	0E	2CCC	97	00	00	5E
2C38	00	31	09	00	2CD0	00	00	90	00
2C3C	7E	09	00	40	2CD4	00	45	13	00
2C40	09	00	5B	09	2CD8	9B	13	00	36
2C44	00	45	13	09	2CDC	00	00	D3	00
2C48	72	00	00	6D	2CE0	00	4F	00	00
2C4C	09	00	DB	00	2CE4	7A	09	00	3B
2C50	00	86	09	00	2CE8	00	00	92	00
2C54	BB	09	00	86	2CEC	00	86	00	00
2C58	09	00	B5	09	2CF0	DE	00	00	36
2C5C	00	4A	13	00	2CF4	13	00	86	00
2C60	A5	09	00	72	2CF8	00	40	13	00
2C64	00	00	CD	00	2CFC	87	00	00	45
2C68	00	6D	31	00	2D00	13	00	84	00
2C6C	C4	00	00	4F	2D04	00	40	00	00
2C70	00	00	6F	09	2D08	A7	00	00	40
2C74	00	54	13	00	2D0C	13	00	EA	00
2C78	80	13	00	59	2D10	00	45	09	00
2C7C	27	00	B6	09	2D14	97	00	00	4A
2C80	00	45	09	00	2D18	09	00	A9	00
2C84	9B	00	00	8B	2D1C	00	54	1D	00
2C88	00	00	CD	00	2D20	9C	09	00	68
2C8C	00	54	40	00	2D24	09	00	D5	00
2C90	AB	00	00	45	2D28	00	68	13	09
2C94	4A	00	B1	00	2D2C	11	52	19	50

Table 8. Parameters for Converting Intelligibility Measure to Intelligibility Score

	A_H	A_L	A_H	A_L
2E00	09	04	05	03
2E04	0A	04	0A	04
2E08	05	03	04	02
2E0C	05	02	07	03
2E10	09	01	08	01
2E14	09	02	09	04
2E18	09	02	09	03
2E1C	08	02	08	02
2E20	04	01	0A	03
2E24	09	04	0A	03
2E28	08	05	09	04
2E2C	09	02	07	02
2E30	04	03	07	03
2E34	04	02	06	03
2E38	09	04	07	02
2E3C	04	02	09	05
2E40	03	02	04	01
2E44	06	02	06	03
2E48	09	05	06	03
2E4C	07	02	06	03
2E50	09	01	06	01
2E54	06	03	05	03
2E58	08	04	07	03
2E5C	04	02	0A	06
2E60	05	02	09	02
2E64	0A	04	08	01
2E68	09	02	04	02
2E6C	06	04	07	02
2E70	0A	07	08	01
2E74	05	02	05	03
2E78	0A	03	08	02
2E7C	06	04	09	04
2E80	06	02	07	05
2E84	09	03	09	03
2E88	09	07	07	02
2E8C	06	04	05	04
2E90	05	04	08	03
2E94	06	05	07	04

	A_H	A_L	A_H	A_L
2E98	09	04	0A	03
2E9C	07	02	09	03
2EA0	0A	08	0A	03
2EA4	08	05	06	03
2EA8	05	03	07	03
2EAC	07	02	0A	04
2EB0	08	06	08	05
2EB4	08	05	08	03
2EB8	07	04	08	03
2EBC	0A	03	0A	03
2EC0	04	02	06	02
2EC4	07	05	04	02

3. TEST RESULTS

In order to permit calibration and evaluation of the system, intelligibility tests were conducted with human subjects. The word list used, which is the same as that recorded on the test tape, is shown in Table 9.

From the master tape, additional tapes were prepared with increasing noise. These tapes were intermixed with other word tapes with varying amounts of noise in order to reduce the likelyhood of word memoriza-by the human subjects. To further reduce the likelyhood of word memorization the noisier tapes were played first during the intelligibility tests.

A summary of the intelligibility test results is shown in Table 10. The reference S/N shown in Table 10 corresponds approximately to a peak voice signal/noise spectral density of 63 dbHz. Because of the difficulty of defining accurately the signal power in a voice signal all values in Table 10 are shown relative to a reference.

In addition to the overall results shown in Table 10, detailed results were obtained for the starting and ending phonemes of individual words in order to permit the determination of the parameters used for conversion in the system from intelligibility measure to intelligibility score.

The same tapes used for obtaining intelligibility scores with human listeners were then run through the system and intelligibility scores were obtained. These are tabulated in Table 11.

A graph showing intelligibility scores as a function of signal to noise ratio for both human listeners and the

Table 9. Word List

click	brass
gob	eye
slice	slush
pack	ace
rouge	cart
rap	in
flash	pad
route	quip
salve	cork
pew	did
theme	crate
wretch	skid
wash	fair
web	threw
clog	robe
soak	get
seed	joke
wise	duke
hump	lid
walk	gang
beard	puss
tilt	base
judge	roost
mow	souse
sigh	fast

Table 10. Number of Missed Words as Function of Signal to Noise Ratio

	No Noise	Ref. S/N	-3db	-6db	-9db	-12 db	-15 db	-18 db	-21 db
Listener #1	2	12	17	20	24	29	35	38	37
2	4	14	15	20	24	31	27	28	29
3	3	11	11	17	20	20	27	24	33
4	2	15	13	18	20	24	26	32	31
5	7	17	17	22	23	30	33	30	34
6	6	14	17	18	25	24	29	27	28
7	4	13	15	16	19	31	31	34	31
8	6	12	14	24	21	27	30	32	36
9	7	12	15	16	20	26	28	29	30
Average	4.6	13.3	14.9	19	19.6	26.9	29.6	30.4	32.1
Int. Score (%)	91	73	70	62	61	46	41	39	36

Table 11. Intelligibility Scores Obtained with Intelligibility Scoring System

	No Noise	Ref. S/N	-3db	-6db	-9db	-12 db	-15 db	-18 db	-21 db
Int. Score (%)	99	70	62	54	44	36	33	29	26

intelligibility scoring system is presented in Figure 7. The intelligibility scoring system gives a score of 99 (almost 100) with no noise while the listener score is 91 because human listeners will misinterpret some words even with an ideal communication system. The intelligibility scoring system also gives a smoother drop-off with decreasing signal to noise ratio.

The close match between the results obtained with the intelligibility scoring system and those obtained with human listeners is not surprising, because the program parameters used in the scoring system were selected to provide such a close match. In general, the program parameters can be adjusted to provide almost any characteristics desired.

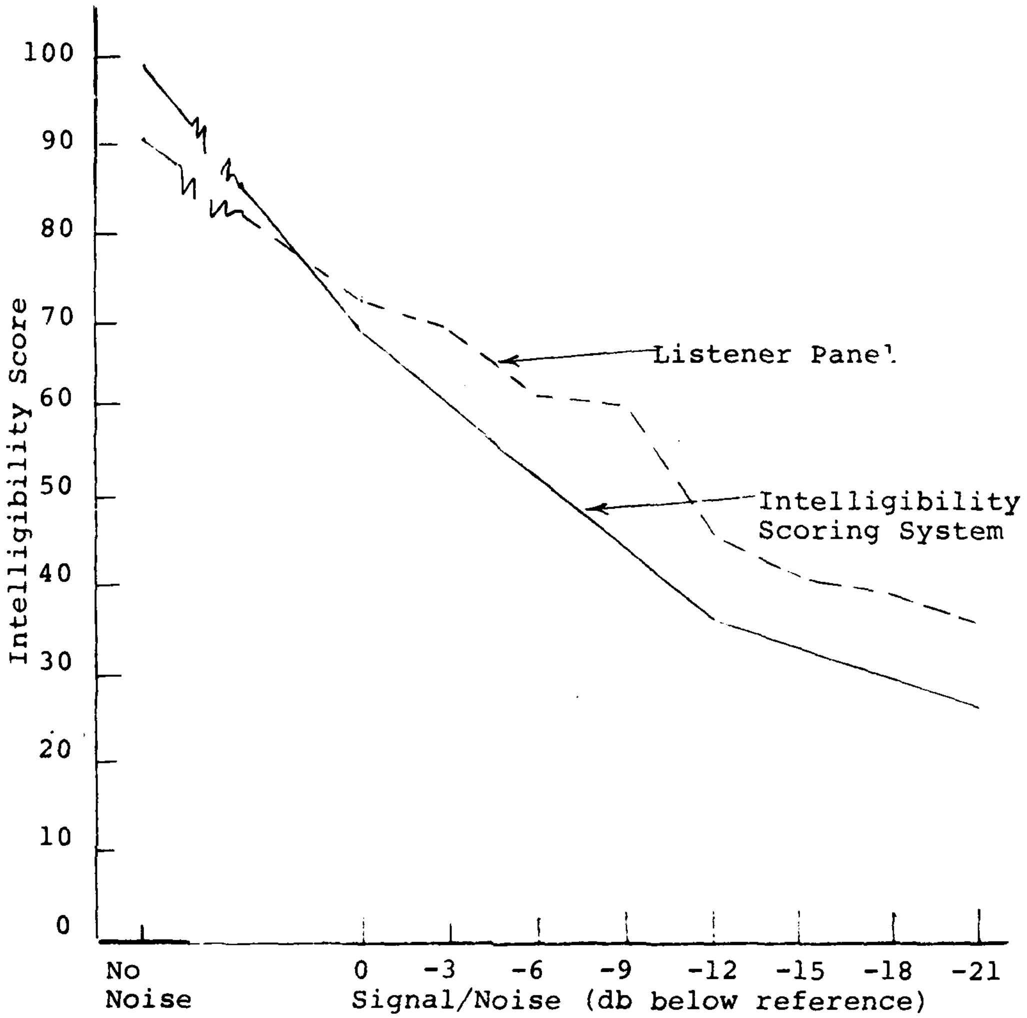

Figure 7. Intelligibility Scores for Intelligibility Scoring System and Listener Panel

4. RECOMMENDATIONS

The tests described in Section 3 were for a linear system in which noise was simply added. Additional tests should be conducted with non-linear systems such as CVSD in order to establish the usefulness of the system.

The system, as designed, can give erroneous results if the tape recorders used do not have very accurate speed control. The time from the timing code to the first sample measurement is a function of tape speed. Some provision should be made for measuring tape speed through the use of multiple timing codes or by other methods, and using this measurement to make software adjustments in measurement times.

As described, the system must decode every timing code in order to time to the sample measurements. It also counts the timing codes in order to keep track of which word is currently being analyzed. Should a timing code be missed, the count would be in error, and the resultant score would be meaningless. Since there is, in any real system, some finite probability that a timing code will be missed, some provision should be made to estimate the time of occurrance of a timing code if one is missed.

REFERENCES

(1) "Voice Coding and Intelligibility Testing for a Satellite Based Air Traffic Control System", by J. N. Birch and N. R. Getzin, The Magnavox Company, April 1971.

(2) "A Comparison of Voice Communication Techniques for Aeronautical and Marine Applications", S.J. Campanella and J. A. Sciulli; COMSAT Technical Review, Vol.2, No. 1, Spring 1972.

(3) "Methods for the Calculation of the Articulation Index", American National Standards Institute, Inc. Approved January 6, 1969.

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.


3