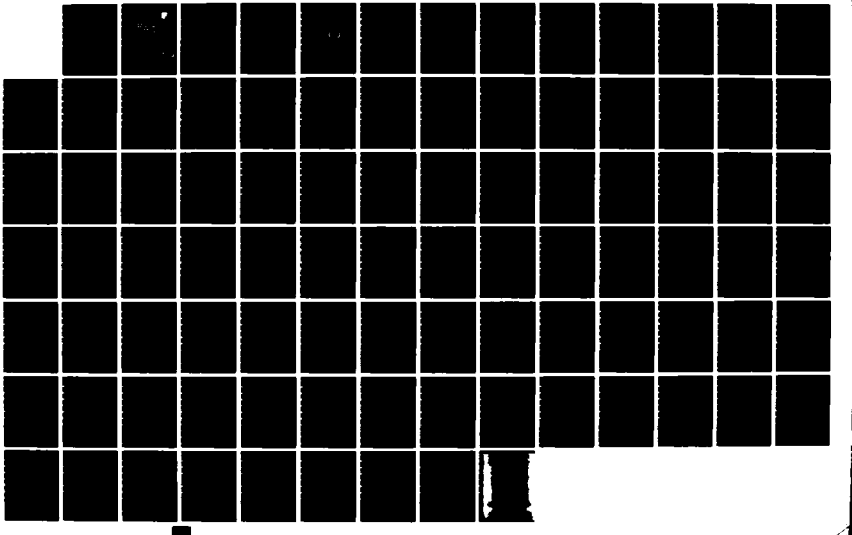
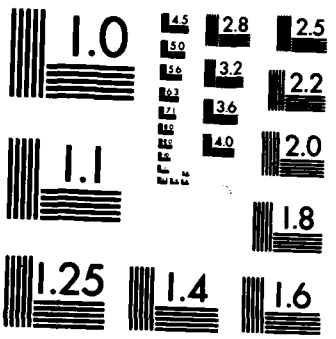
AD-A141 333 NARROWBAND (LPC-10) VOCODER PERFORMANCE UNDER COMBINED EFFECTS OF RANDOM.. (U) ROME AIR DEVELOPMENT CENTER GRIFFISS AFB NY C P SMITH DEC 83 RADC-TR-83-293 1/1

UNCLASSIFIED F/G 17/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

RADC-TR-83-293
In-House Report
December 1983

NARROWBAND (LPC-10) VOCODER PERFORMANCE UNDER COMBINED EFFECTS OF RANDOM BIT ERRORS AND JET AIRCRAFT CABIN NOISE

Caldwell P. Smith

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

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

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-83-293 has been reviewed and is approved for publication.

APPROVED:

J. P. VETRANO, Chief
COMSEC Engineering Office
Electromagnetic Sciences Division

APPROVED:

ALLAN C. SCHELL, Chief
Electromagnetic Sciences Division

FOR THE COMMANDER:

JOHN A. RITZ
Acting Chief, Plans Office

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (EEV) Hanscom AFB MA 01731. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-83-293	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) NARROWBAND (LPC-10) VOCODER PERFORMANCE UNDER COMBINED EFFECTS OF RANDOM BIT ERRORS AND JET AIRCRAFT CABIN NOISE		5. TYPE OF REPORT & PERIOD COVERED In House. June 82 - Sept. 83 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Caldwell P. Smith		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Rome Air Development Center (EEV) Hanscom AFB Massachusetts 01731		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 33401F 12050201
11. CONTROLLING OFFICE NAME AND ADDRESS Rome Air Development Center (EEV) Hanscom AFB Massachusetts 01731		12. REPORT DATE December 1983 13. NUMBER OF PAGES 90
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

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)

18. SUPPLEMENTARY NOTES

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

Vocoder
Speech intelligibility
Digital voice communication
Intelligibility

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

Speech intelligibility and voice quality tests were conducted to evaluate the performance of the DOD standard LPC-10 narrowband vocoder algorithm under degradation caused by the combined effects of random bit errors in the data path and background noise accompanying the input speech, as occurs in the acoustic environment in the cabins of an E4B aircraft during operational flight.

Vocoder intelligibility scores and voice quality scores indicated that the in-flight acoustic noise environments of the E4B Briefing Room, Battlestaff

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

Unclassified

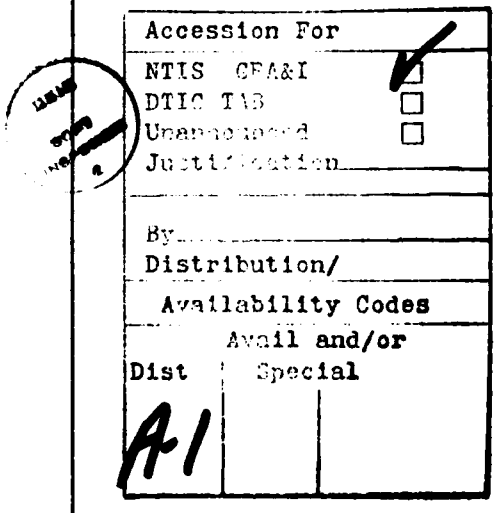
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (contd)

Compartment, and NCA Compartment were alike in their effects on overall vocoder performance. Composite performance data pooled for the three compartments indicated that vocoder intelligibility averaged 80.2 ("fair") with no bit errors, and fell 4-1/2 points with each percent of bit errors incurred in the data path. Thus, the average intelligibility fell to 77.9 ("poor") at 1/2 percent error rate, to 73.4 ("very poor") at 1-1/2 percent error rate, and to 68.9 ("unacceptable") at 2-1/2 percent error rate.

Voice quality scores averaged 43.3 ("fair") in the absence of data errors and fell 2.3 points with each percent of incurred bit errors. Thus, at 1 percent error rate, the average quality score was 41.0 ("poor") and dropped to 35.2 ("very poor") at 3-1/2 percent error rate. These estimates were obtained from linear regression modeling of the relationship between performance (intelligibility scores and voice quality ratings) and bit error rate.

Similar intelligibility performance was obtained for the acoustic environment of an EC-135 aircraft, but lower speech-quality scores; this finding was probably a result of using a pressure microphone (not noise cancelling) in that environment.



Contents

Illustrations

Tables

Narrowband (LPC-10) Vocoder Performance Under Combined Effects of Random Bit Errors and Jet Aircraft Cabin Noise

1. INTRODUCTION

The performance of digital voice communications processors (vocoders, etc) over degraded communications channels, such as channels where data errors are incurred because of interference, power or bandwidth limitations, or other causes, is of prime importance in designing communications systems. Therefore, tests to determine the suitability of a voice processor for tactical or strategic communications usually include speech intelligibility and voice quality evaluations under impairments caused by bit errors in the data path.

1.1 Linear Regression Modeling

It has been determined that speech intelligibility scores and voice quality scores from tests of vocoders and wideband voice processors under random bit error conditions tend to follow a linear relationship with bit error rate.[1,2] The

---

(Received for publication 22 December 1983)

1. Smith, C. P. (1977) Intelligibility Performance of Narrowband Linear Predictive Vocoders in the Presence of Bit Errors, AF Electronic Systems Division, Hanscom AFB, MA 01731, ESD-TR-77-328, AD A051323.
2. Smith, C. P. (1983) Relating the performance of speech processors to the bit error rate, Speech Tech. 2(No. 1).

microphones and the particular male and female speakers used for testing are factors that can impart an overall bias to scores. However, assuming that tests have been conducted at several bit-error rates with these factors held constant, voluminous test data can be reduced to simple linear expressions of the form

$$\hat{S} = A + Bx$$

where $\hat{S}$ provides an estimate of average speech intelligibility score (or voice quality score); x = bit error rate in percent; and the regression coefficients A and B express the origin and slope, respectively, of a linear function estimating performance vs bit error rate that is fitted to the data values. The coefficient A expresses the expected score with no bit errors, the origin of the regression line. The coefficient B expresses the slope of the line, the amount that the score will drop with each percent bit error rate. For example, confidence limits calculated from the data values can establish the 95 percent probability bounds of the predicted average score at any error rate, and the confidence bounds for the data population, that is, the collection of intelligibility scores or quality scores from which the regression model was calculated.[3]

1.2 Combined Bit-Error and Acoustic Environment Effects

In earlier studies, extensive tests of vocoders and wideband speech processors [continuous variable-slope delta modulation, or CVSD, at 16 kilobits per second (K bps) and 32 K bps data rates] were conducted under random and block errors with "ideal" input speech signals, that is, speech from a quiet environment and through a high quality microphone. The study reported here evaluated the combined effects of random errors imposed on the vocoder data and of noise combined with the input speech, noise representing the acoustic environments found in the cabins of jet aircraft during flight. The voice processor tested was the Department of Defense (DOD) standard LPC-10 narrowband vocoder algorithm.[4] Acoustic noise environments were those measured in E4B and EC-135 aircraft during flight. Microphones used for the tests were the resident microphones, the handset instruments installed in these aircraft. Performance data from earlier evaluations of voice processor performance under bit error impairments with speakers in a quiet environment are also presented in this report for comparison.

---

3. Snedecor, G.W., and Cochran, W.G. (1967) Statistical Methods, Iowa State University Press, Ames, Iowa.

4. Tremain, T. (1982) The government standard linear predictive coding algorithm: LPC-10, Speech Tech. 1(No. 2).

1.3 Background

This study was conducted by the COMSEC Engineering Office (RADC/EEV), Electromagnetic Sciences Division, RADC, Hanscom AFB, Mass., under sponsorship of the AF Electronic Systems Division (ESD/SCS). The work involved a team effort, with contractor participation provided by Ketron, Inc., Wayne, Pa., and by Dynastat, Inc., Austin, Tex. Mr. Charles Teacher, Ketron's principal investigator, made significant contributions by planning and participating in the operational flights. He had the lead role in taking measurements aboard the aircraft, collecting data, and making live speech recordings. Mr. Teacher also analyzed the acoustic environments and microphone characteristics.[5]

Dynastat furnished the standard speakers used for live speech recordings aboard the aircraft and prepared master tapes of simulations of the acoustic environments. Here, "standard" speakers refers to individuals whose voices are recorded in an extensive DOD tape library of speech materials for intelligibility and voice quality testing, and on whom extensive data has been collected from evaluations of the performance of DOD and industry voice processors under many test conditions. Dynastat conducted tests with listener crews to determine speech intelligibility and voice quality, using government-furnished recordings prepared at the COMSEC Engineering Office, RADC/EEV while testing the DOD standard vocoder.

RADC/EEV provided overall management and direction of the effort as well as the data analyses and interpretations reported here.

2. TASKS

Accomplishment of this work required the following sequence of tasks:

2.1 Characterization of the Acoustic Environments Aboard the Aircraft

HQ SAC allowed participation in flights from Offut AFB, Nebr., in June and July 1982. During those flights on E4B and EC-135 aircraft, the following tasks were accomplished:

(1) Sound pressure levels were measured at several positions in each aircraft.

(2) Live recordings of speech intelligibility test materials and voice quality test materials were prepared, using as speakers Mr. Teacher from Ketron and speakers provided by Dynastat.

(3) Acoustic background noise in the cabins of the aircraft was recorded with

---

5. Teacher, C.F. (1983) Secure Voice Evaluations: Characterization of Aircraft Noise and Audio Systems in the EC-135 and E4B Aircraft, Report KET-2709-2, Ketron, Inc., Wayne, Pa.

a portable stereo tape recorder for subsequent uses in simulating the acoustic environments.

2.2 Preparation of Master Tapes for Speech Tests and Evaluations

The number of live speech test recordings aboard the aircraft was extremely limited. Therefore, Dynastat expanded speech intelligibility and voice quality test materials representative of the E4B Battlestaff Compartment, Briefing Room, and NCA Compartment, and of the EC-135 Battlestaff Compartment and Radio Compartment acoustic environments by preparing simulations. Acoustic environments were reproduced in a sound room by reproducing the in-flight recordings of cabin noise, with the sound pressure level of the reproduced cabin noise adjusted to match the levels measured aboard the aircraft. The stereo noise tapes were reproduced with two loudspeakers placed in corners of the room. A measurement and calibration procedure verified that the reproduced noise field was relatively uniform in the vicinity of the speaker subjects and their microphones, and that the sound pressure levels of the simulations reproduced those measured aboard the aircraft. Normal sidetone signals were furnished for the speaker subjects during the recordings. Master tapes prepared in these simulations are listed in Appendix B.

2.2.1 MICROPHONES

Microphones were mounted in handsets and clamped in supports keeping them at the optimum placement from the speaker's lips.

For recordings in the E4B environment, a Roanwell Model 240100001 pressure-gradient (noise cancelling) microphone mounted in a handset was used. This microphone is used for command and control functions on the E4B; it is referred to in this report as the "E4B resident microphone."

For the EC-135 acoustic environment test, a Roanwell Model 60150 pressure microphone mounted in a handset, the command and control microphone used on that aircraft, was used. The Model 60150 is not a noise cancelling microphone, and, while test results indicated that speech intelligibility of the vocoder was similar in the two aircraft, the voice quality ratings obtained in the EC-135 acoustic environment were poorer than those in the E4B, probably because the speech-to-noise ratio was poorer.

2.2.2 SPEAKER SUBJECTS

Six male and three female speakers were speaker subjects in the simulations of the acoustic environments; of these, three of the male speakers were common to the live recordings aboard the aircraft. Four "scramblings" (randomizations) of

the Diagnostic Rhyme Test (DRT)[6, 7, 8] intelligibility word lists were prepared for each speaker/test-condition combination. The design of the DRT intelligibility test makes it unlikely that subjects will memorize word lists with repeated presentations; for these studies, the many equivalent versions of word lists guaranteed that memorization would not occur.

2.3 Vocoder Tests

Vocoder tests were carried out at RADC/EEV's speech test and evaluation facility at Hanscom AFB, Mass. The facility has been used for voice processor tests for the DOD Digital Voice Processor Consortium over the past 3 years. It is equipped with professional tape recorders, random bit-error generators, a hardware vocoder that incorporates the DOD standard LPC-10 narrowband vocoder algorithm, wideband 16K bps and 32K bps CVSD voice processors, and the DOD library of master tape recordings prepared specifically for testing speech processors for intelligibility and quality. Bit error tests were conducted with random errors at 0-, 1/2-, 1-, 2-, and 5-percent error rates, with replication of the test at each error rate using different scramblings of the DRT word lists.

2.4 Listener Tests

For listener tests, tape recordings of the vocoder speech were prepared at RADC/EEV and forwarded to Dynastat for presentation to listener crews of eight or more subjects. The speech signals were presented *diotically (to both ears)* over headphones, with subjects assembled in an acoustically treated room. Listener tests were "blind": neither the listeners nor test administrators knew the identity of processors or test conditions under evaluation. Test materials were identified only by the assigned serial numbers, a procedure routinely used to guarantee the integrity of test results.

2.5 Analysis of Performance Data

Diagnostic scores and overall scores were determined from listener responses by Dynastat. Listings of diagnostic scores and voice quality ratings with standard errors were supplied to RADC.

---

6. Voiers, W.D.; Cohen, M.E.; and Mickunas, J. (1965) Evaluation of Speech Processing Devices: 1. Intelligibility, Quality and Speaker Recognizability, AFCRL-65-826, Final Report, Contract AF19(628)-4195, AD 627320.
7. Voiers, W.D. (1977) Diagnostic evaluation of speech intelligibility, in Speech Intelligibility and Speaker Recognition, 2, M.E. Hawley, Ed., Benchmark Papers in Acoustics, Dowden, Hutchinson and Ross, Stroudsburg, Pa.
8. Voiers, W.D. (1983) Evaluating processed speech using the Diagnostic Rhyme Test, Speech Tech. 1(No. 3).

The performance data were analyzed at RADC for the following assessments:

(1) Similarities of acoustic environments in the aircraft compartments;

(2) Comparisons of live and simulated environments;

(3) Computation of regression models expressing the relationships between performance and error rate; and

(4) Comparisons of acoustic environments in the aircraft with a quiet environment.

3. APPROACH

As vocoder equipments have become more sophisticated and complex, testing procedures have become refined, and informal listening judgments by "experts" have been replaced by standardized multiple-speaker diagnostic tests presented to listener crews of eight or more subjects. After frustrating experiences with older channel-vocoders, we appreciated that good speech intelligibility was a necessary but not sufficient attribute for measuring user acceptance. Speech quality tests were developed to supplement intelligibility tests,* and these, in turn, have been refined to provide diagnostic information on speech quality.[9]

3.1 The Diagnostic Rhyme Test (DRT)

Over the years, vocoder intelligibility has been evaluated with phonetically balanced (PB) word lists, the Fairbanks Rhyme Test,[10] the Modified Rhyme Test (MRT),[11] the Consonant Rhyme Test (CRT),[12] and the Diagnostic Rhyme Test (DRT).[6,7,8] Of these, the most widely used in the past decade has been the DRT of Voiers, Cohen, and Mickunas,[6] which has become the preferred intelligibility test method of the DOD Digital Voice Processor Consortium. Numerous tests of voice processors have shown that any processor that obtains a high intelligibility score on multiple-speaker tests by this method will be capable of providing highly

---

*Speech quality tests supplement, rather than replace, intelligibility tests because high "quality" scores do not necessarily imply high intelligibility. For example, speech low-pass filtered at 2 kHz results in high quality scores even though intelligibility is significantly impaired.

9. Voiers, W.D. (1977) Diagnostic acceptability measure for speech communications systems, IEEE Proc. ICASSP, 77CH1197-3 ASSP, 204-207.

10. Fairbanks, G. (1958) Test of phonemic differentiation: the Rhyme Test, J. Acoust. Soc. Am. 30:596-600.

11. House, A.S.; Williams, C.E.; Hecker, M.H.L.; and Kryter, K.D. (1965) Articulation testing methods: Consonantal differentiation with a closed response set, J. Acoust. Soc. Am. 37:158-166.

12. Preusse, J.W. (1959) The Consonant Recognition Test, U.S. Army Electronics Command, ECOM-3205, Fort Monmouth, N.J.

intelligible speech communications for ordinary conversational speech, and will provide even higher margins of performance when messages are highly stereotyped, a condition typical of military communications, where context and redundancy improve intelligibility.

3.2 Categories of DRT Intelligibility Scores

The DOD Digital Voice Processor Consortium has now established a scale of descriptive categories for DRT intelligibility scores. The categories describe the level of intelligibility associated with different ranges of intelligibility scores and are illustrated in Table 1. Scores of 96 and higher are categorized as "excellent." Such scores are obtained with high quality speech under quiet conditions, such as the DRT master tapes used in evaluating "ideal" conditions. At the other extreme, scores below 70 are categorized as "unacceptable." An example of unacceptability is vocoder intelligibility with a speaker in the noise environment of a military helicopter. Between these extremes are categories "very good" for scores from 96 to 91; "good" for scores from 91 to 87, obtained with typical commercial telephony within the continental United States; "moderate" for scores from 87 to 83; "fair" for the range from 83 to 79; and "poor" for scores from 79 to 75. Intelligibility scores between 75 and 70 are categorized as "very poor."

These values represent a consensus of the DOD Digital Voice Processor Consortium Test and Evaluation Committee. Further work is required to validate these values with judgments of operational personnel.

3.3 Voice Quality Tests: The Diagnostic Acceptability Measure (DAM) Test

Each vocoder test involved paired intelligibility and voice quality tests using identical speakers and conditions. Speech samples for evaluating quality used connected speech rather than words spoken in isolation. In the Diagnostic Acceptability Measure (DAM) test of Voiers,[9] the processed speech samples were presented to listener crews who made judgments of quality attributes of the speech signal and the background noise, caused by the combination of bit errors and acoustic noise environment, that accompanied the signal.

Dynastat's analyses of listener responses were incorporated into data summaries that included scores for speech signal quality, designated Composite Signal Acceptability (DAM/CAS) scores; background signal quality, designated Composite Background Acceptability (DAM/CBA) scores; and overall composite quality, designated Composite Acceptability Estimate (DAM/CAE) score, a weighted combination of the CSA and CBA scores. (Overall quality can be estimated from the DAM/CSA and DAM/CBA scores by expressing the scores as fractions and multiplying the CSA and CBA scores. For example, with a CSA score of 80 and a CBA

Table 1. Categories of DRT Intelligibility Scores With Examples Based on Typical DRT Scores for Male Speakers

DRT Score	Category	Examples	Qualifiers for these examples
100	Excellent	Unfiltered speech Speech low-pass filtered at 4 kHz	Speech from a quiet environment; no significant distortions; high-quality microphone
96	Very Good	CVSD at 32 K bps CVSD at 16 K bps	Error rate less than 1%; speech from a quiet environment
91	Good	Typical commercial telephony within continental USA APC Processor at 9600 bps LPC-10 Vocoder at 2400 bps, no bit errors	Speech from a quiet environment
87	Moderate	LPC-10 Vocoder with bit error protection, at 2400 bps with 2% random bit errors	Speech from a quiet environment
83	Fair	LPC-10 Vocoder without bit error protection, at 2400 bps with 2% random bit errors	Speech from a quiet environment
79	Poor	LPC-10 Vocoder with bit error protection, at 2400 bps with 5% random bit errors	Speech from a quiet environment
75	Very Poor	Experimental 800 bps voice processor with no bit errors	Speech from a quiet environment
70	Unacceptable	LPC-10 Vocoder at 2400 bps	Speech from a helicopter noise environment

score of 70, multiplying 0.80 by 0.70 results in 0.56, or a DAM/CAE score of approximately 56.)

3.4 Categories of DAM/CAE Voice Quality Scores

Interpretation of voice quality scores is more difficult than interpretation of intelligibility scores. Quality scores do not cover a range from 0 to 100, but typically occur over a more restricted range from about 20 to 70. A DAM/CAE score

of 20 represents extremely distorted and noisy speech (a score below 30 generally implies unacceptable speech quality). A DAM/CAE score of 70 represents speech quality approaching "high fidelity."

Our experience in interpreting numerous tests has resulted in assigning the descriptive categories for quality scores shown in Table 2. Voice quality (DAM/-CAE) scores above 64 are categorized as "excellent," and scores from 64 to 58 are "very good"; scores in the 64-58 range have been obtained for a 32 K bps CVSD voice processor with no bit errors, and input speech from a quiet environment and a high quality microphone. Scores for the same processor drop to the "moderate" range, between 53 and 48, when the speech is from an "office noise" environment. Quality scores in the range from 48 to 42 are categorized as "fair." Examples are the scores obtained for a 16 K bps CVSD processor at 5 percent random bit

Table 2. Categories of DAM/CAE Voice Quality Scores With Examples Based on Typical DAM/CAE Scores for Male Speakers

DAM/CAE Score	Category	Examples	Qualifiers for these examples
	Excellent	High fidelity speech	From a quiet environment
64			
	Very Good	CVSD at 32 K bps, with no bit errors	Speech from a quiet environment
58			
	Good	CVSD at 16 K bps, with no bit errors	Speech from a quiet environment
53			
	Moderate	CVSD at 16K bps, with no bit errors	Speech from an office noise environment
48			
	Fair	CVSD at 16K bps, with 5% random bit errors LPC-10 Vocoder with bit error protection, at 2400 bps with 1% random bit errors	Speech from a quiet environment
42			
	Poor	LPC-10 Vocoder with bit error protection, at 2400 bps with 2% random bit errors	Speech from a quiet environment
36			
	Very Poor	Experimental 800 bps voice processor, with no bit errors	Speech from a quiet environment
30			
	Unacceptable	LPC-10 Vocoder at 2400 bps	Speech from a helicopter noise environment

errors, or the DOD standard LPC-10 narrowband vocoder with 1 percent random bit errors, with speech from a quiet acoustic environment in both examples. Between 42 and 36, quality scores are categorized as "poor." One example is the quality (DAM/CAE) score obtained for the DOD standard LPC-10 vocoder at a 2 percent random bit error rate, where the input speech is from a quiet environment and through a high quality dynamic microphone. The range of quality scores from 36 to 30 is categorized as "very poor." An example is an experimental voice processor operating error-free at 800 bps with a high-quality input speech signal. Voice quality scores below 30 are considered "unacceptable"; scores in this range have been obtained from tests of the DOD standard LPC-10 vocoder where the input speech, at a very poor signal-to-noise ratio, came from a military helicopter noise environment.

3.5 Discussion of Categories of Scores

Descriptive categories of intelligibility and voice quality scores presented in Tables 1 and 2 are proposed as an aid in interpreting evaluation data. The designated scales are tentative and may undergo revisions after further research and experience. For example, separate scales for categories may be appropriate for distinguishing between levels of performance associated with tactical communications and those associated with strategic and executive communications. Speech quality levels rated poor for executive communications may be rated good by communicators in the tactical world. Additional research is required to clarify this question.

4. CHARACTERIZATIONS OF AIRCRAFT NOISE ENVIRONMENTS

Teacher[5] has reported the results of measuring the background noise levels aboard the E4B and EC-135 aircraft; the measured sound pressure levels are summarized in Table 3. The background noise levels measured in the E4B Battlestaff

Table 3. Sound Pressure Levels Measured Aboard E4B and EC-135 Aircraft

E4B Battlestaff Compartment	88 dbc
E4B NCA Compartment	78 dbc
E4B Briefing Room	84 dbc
EC-135 Battlestaff Compartment	89 dbc
EC-135 Radio Compartment	88.5 dbc

Compartment and Briefing Room closely match the noise levels measured in an operational flight of an E3A aircraft. Simulations of the acoustic environments prepared by Dynastat are described in Section 2.2.

5. VOCODER TESTS

5.1 Vocoder Test Setup

The setup for vocoder testing is shown in Figure 1. The live speech recorded

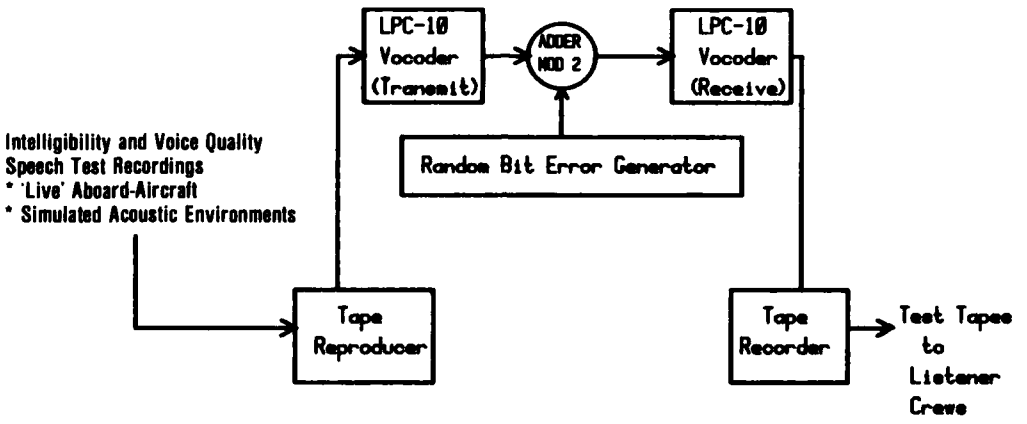


Figure 1. Block Diagram of Vocoder Test Setup

aboard the aircraft and master tapes prepared by simulating the acoustic environments were used as speech input materials to the vocoder. A random bit-error generator created errors in the vocoder data path at the specified error rates. Output speech from the receiving vocoder was recorded for subsequent evaluations by listener crews. With every bit-error condition, the test was repeated a second time with different randomizations of the DRT word lists. Table 4 lists the vocoder test conditions and master tapes used for each condition.

5.2 Live vs Simulated Tests

It has been argued that testing procedures using live speech input are superior to those using speech recordings. The distinction is not clear-cut because several levels of "liveness" must be considered. At one extreme of test authenticity, the vocoder could be taken aboard the aircraft with teams of speaker subjects and lis-

Table 4. Schedule of Master Tapes Processed for Evaluating the DOD LPC-10 Vocoder Performance in the E4B Acoustic Environment

	Zero BER	1/2% BER	1% BER	2% BER	5% BER
Battlestaff	#9540	#9546	#9555	#9563	#9565
Compartment	K10-1.1-A	K10-1.3-A	K10-1.1-A	K10-1.2-A	K10-1.1-A
Speakers:	K10-1.2-B	K10-1.4-B	K10-1.3-B	K10-1.1-B	K10-1.4-B
6 Males	K10-1.3-C	K10-1.1-C	K10-1.3-C	K10-1.4-C	K10-1.1-C
3 Females	#9547	#9541	#9554	#9567	#9562
	K10-1.4-A	K10-1.2-A	K10-1.4-A	K10-1.4-A	K10-1.3-A
	K10-1.1-B	K10-1.3-B	K10-1.2-B	K10-1.2-B	K10-1.4-B
	K10-1.2-C	K10-1.4-C	K10-1.1-C	K10-1.4-C	K10-1.2-C
NCA	#9556	#9548	#9542	#9553	#9566
Compartment	K10-2.3-A	K10-2.3-A	K10-2.1-A	K10-2.4-A	K10-2.3-A
Speakers:	K10-2.3-B	K10-2.4-B	K10-2.2-B	K10-2.2-B	K10-2.1-B
6 Males	K10-2.2-C	K10-2.1-C	K10-2.3-C	K10-2.1-C	K10-2.4-C
3 Females	#9564	#9557	#9549	#9543	#9552
	K10-2.2-A	K10-2.1-A	K10-2.4-A	K10-2.2-A	K10-2.2-A
	K10-2.2-B	K10-2.4-B	K10-2.1-B	K10-2.3-B	K10-2.1-B
	K10-2.1-C	K10-2.4-C	K10-2.2-C	K10-2.4-C	K10-2.3-C
Briefing	#9560	#9568	#9558	#9550	#9544
Room	K10-3.4-A	K10-3.3-A	K10-3.3-A	K10-3.3-A	K10-3.1-A
Speakers:	K10-3.3-B	K10-3.3-B	K10-3.2-B	K10-3.2-B	K10-3.2-B
6 Males	#9569	#9559	#9561	#9545	#9551
	K10-3.1-A	K10-3.2-A	K10-3.1-A	K10-3.2-A	K10-3.4-A
	K10-3.4-B	K10-3.1-B	K10-3.4-B	K10-3.3-B	K10-3.1-B

tener subjects, and tests could be run during a flight. However, not only would this procedure be very costly, but also, inevitably, there would be many uncontrolled variables, including variations in speaker effort, microphone placement, correctness of pronunciation of test words, voice pitch, fluctuations in the acoustic environment, etc. The result would be inflated standard errors and the impaired ability to make close comparisons of test conditions or processor hardware.

At a reduced level of authenticity is the live testing procedure used here: the speaker subjects were recorded aboard the aircraft in the actual acoustic environment, and their tape recordings were then used to test the vocoder. This procedure is much less costly than the first, but is still subject to uncontrolled variations.

At a third level, speaker subjects and listener subjects are placed in simulated acoustic environments, and the tests are conducted live in a sound chamber. With this procedure, greater control can be maintained than with the first two.

A fourth procedure, using recorded speech materials prepared under closely

controlled conditions and validated with testing and comparisons of results, makes it possible to test and retest processors at different times and places without compromising the ability to make close comparisons of results. In addition, testing is less costly because the formidable logistics problems of having speaker subjects and listener subjects, vocoder and test hardware, sound rooms, and test personnel available simultaneously are avoided.

6. INTELLIGIBILITY OF THE DOD LPC-10 VOCODER IN THE E4B ACOUSTIC ENVIRONMENT

Intelligibility scores obtained for male and female speakers from testing the DOD standard LPC-10 narrowband vocoder in the acoustic environments of the E4B Battlestaff Compartment, NCA Compartment, and Briefing Room are presented in Tables A1, A2, and A3. Intelligibility scores obtained with each acoustic environment were analyzed, and linear regression models for the relationship between intelligibility and bit error rate were calculated. The individual speaker's intelligibility scores listed in these tables are presented as scatter diagrams in relation to the computed regression line and 95 percent confidence bounds for the data populations in Figures 2-6.

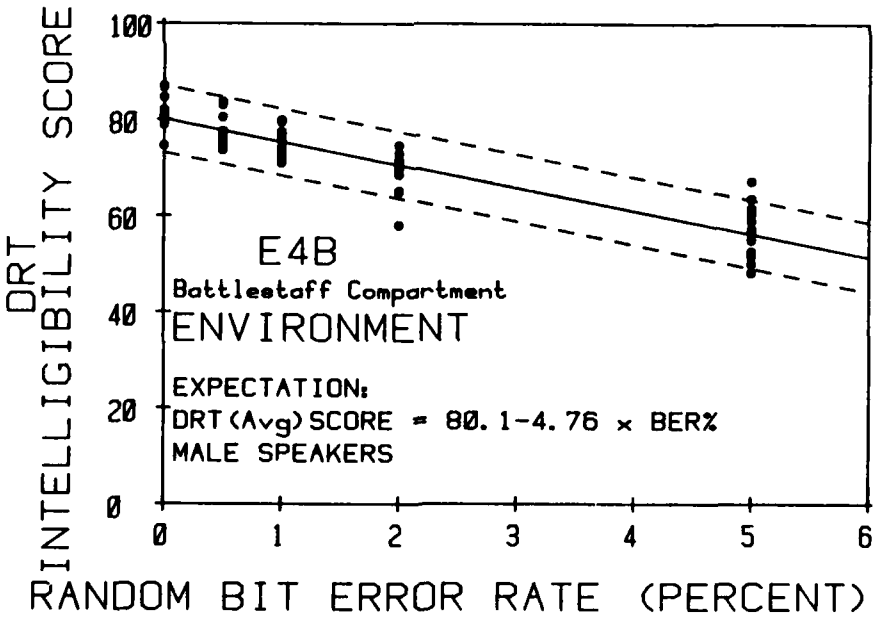


Figure 2. Male Speaker's Intelligibility Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B Battlestaff Compartment Acoustic Environment

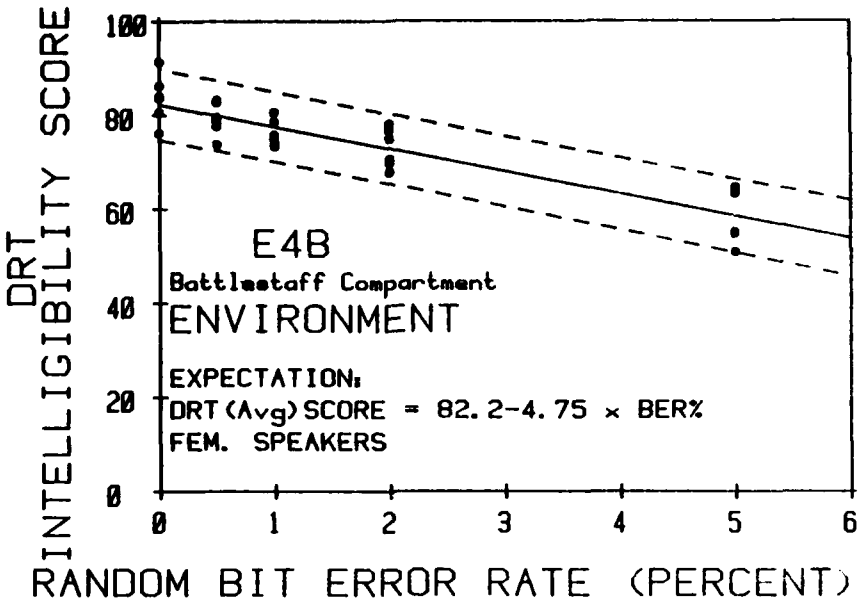


Figure 3. Female Speaker's Intelligibility Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B Battlestaff Compartment Acoustic Environment

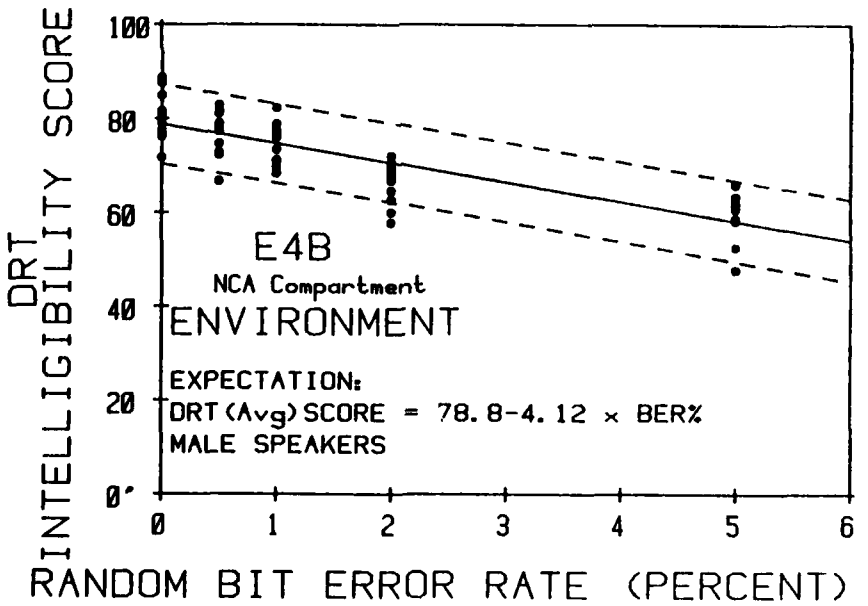


Figure 4. Male Speaker's Intelligibility Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B NCA Compartment Acoustic Environment

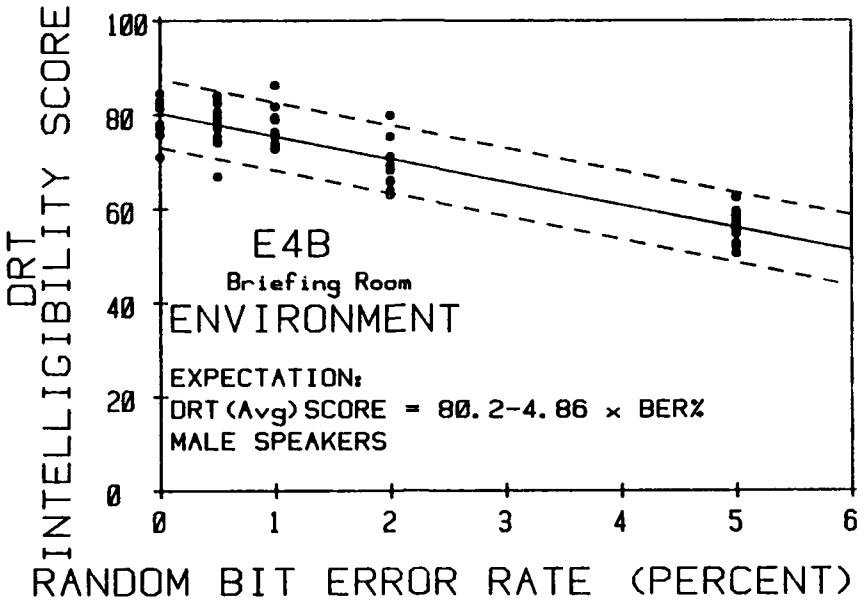


Figure 5. Female Speaker's Intelligibility Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B NCA Compartment Acoustic Environment

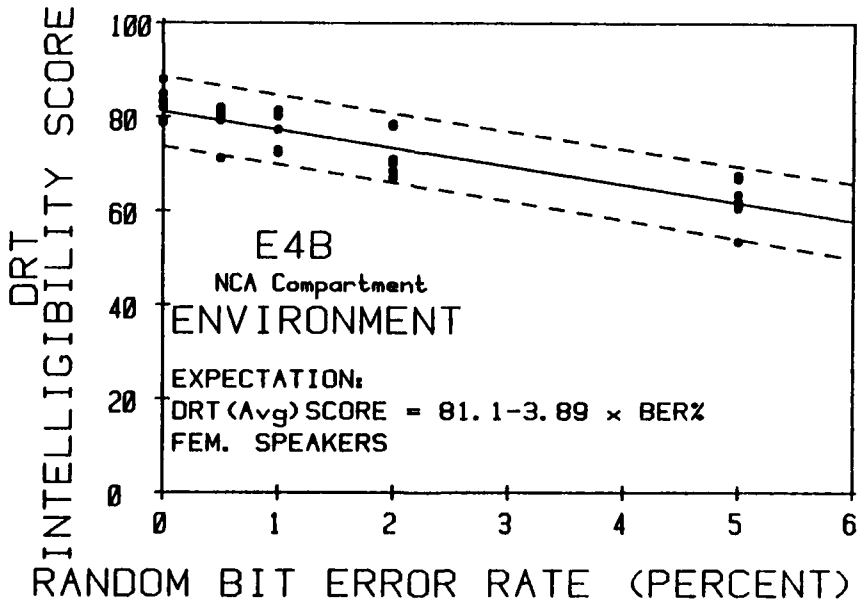


Figure 6. Male Speaker's Intelligibility Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B Briefing Room Acoustic Environment

The acoustic environments were found to be similar in their effects on variation of intelligibility with bit error rate. Regression lines determined for scores from the three environments are compared in Figure 7. The comparison indicates that the lower background noise level of the E4B NCA Compartment may have been involved in producing a small advantage in intelligibility at the higher error rates for speech from that environment. However, the origins of the regression lines are nearly identical, and the advantage is only apparent at error rates where intelligibility is well below 70, a threshold value for minimum acceptable intelligibility. Consequently, the different between regression lines determined for the NCA Compartment and the other two compartments has no practical significance.

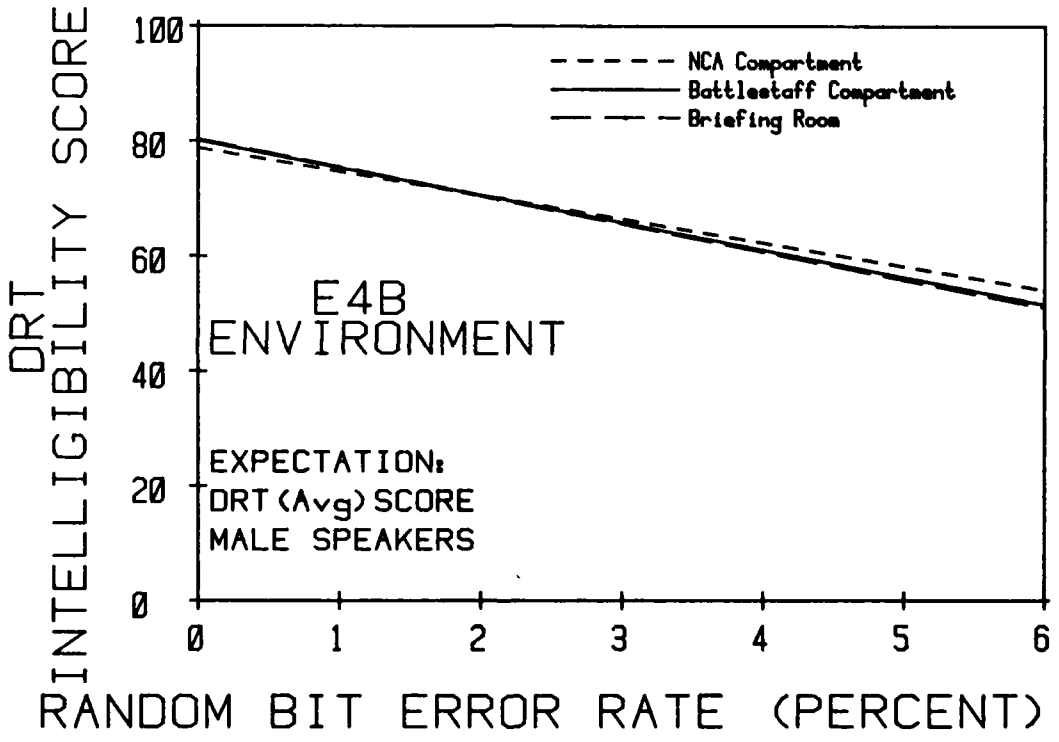


Figure 7. Comparison of Expected Variation in Intelligibility With Bit Error Rate for the DOD LPC-10 Vocoder Tested in the Acoustic Environments of the Three E4B Compartments

6.1 Composite Result for E4B Acoustic Environment: Intelligibility vs Bit Error Rate

Pooling intelligibility scores for the three environments and calculating a re-

gression model for the composite data from male and female speakers, with 95 percent confidence bounds for this data population, resulted in the estimate of overall relationship between intelligibility and bit error rate presented in Figure 8.

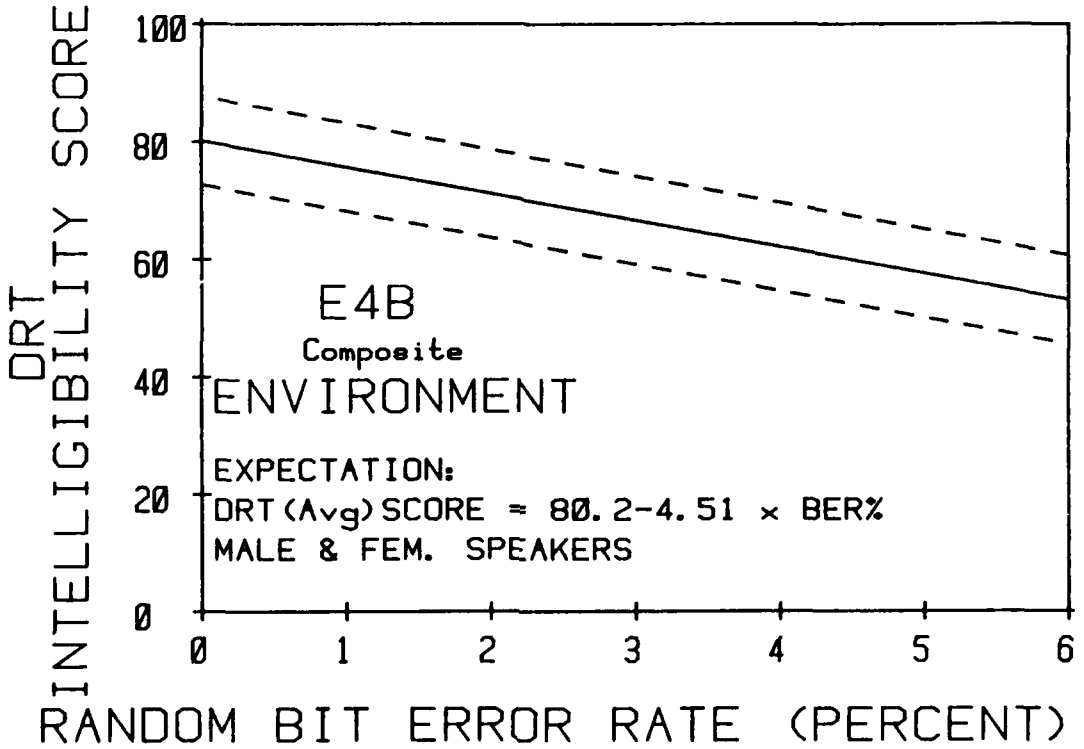


Figure 8. Linear Regression Model Expressing Variation in Intelligibility With Bit Error Rate, With 95% Confidence Limits for Individual Speaker's Intelligibility Scores, for the DOD LPC-10 Vocoder in the Composite E4B Acoustic Environment

The composite data indicated that an average intelligibility score of 80.2 (fair) can be expected with no bit errors, and that bit errors will cause intelligibility to fall 4-1/2 points for each percent bit errors. The regression line indicates that average intelligibility will become unacceptable (will drop below a score of 70) at bit error rates greater than 2.3 percent.

6.2 Intelligibility of Individual Speakers in the Composite E4B Acoustic Environment

Statistical tests of regression models fitted to intelligibility scores obtained for individual speakers in the composite E4B acoustic environment indicated that

slopes of individual speaker's regression lines for vocoder intelligibility in this environment did not differ significantly. However, on the basis of a common slope, differences in the elevations (origins) of individual speaker's regression lines were significant. A new calculation based on a common slope led to the result indicated in Figure 9. The result illustrates how different individuals varied in intelligibi-

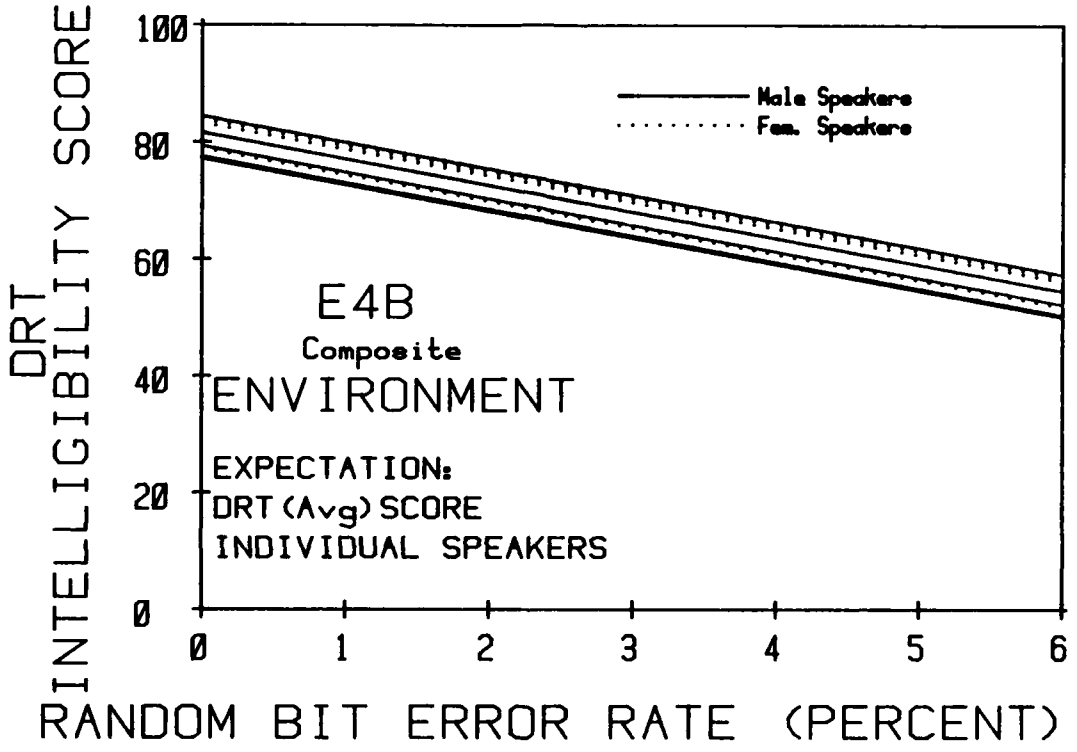


Figure 9. Regression Lines Expressing Variation in Intelligibility With Bit Error Rate of Individual Speakers, in Tests of the DOD LPC-10 Vocoder in the E4B Composite Acoustic Environment

lity under this combination of vocoder, acoustic environment, and bit errors. Origins of regression lines for individual speakers varied from 84.4 for the "best" speaker to 77.2 for the "poorest," with a common slope of -4.51.

7. VOICE QUALITY RATINGS (DAM/CAE SCORES) OBTAINED FOR THE DOD LPC-10 VOCODER IN THE E4B COMPOSITE ACOUSTIC ENVIRONMENT

Overall quality ratings (DAM/CAE scores) from testing the DOD narrowband LPC-10 vocoder in the acoustic environments of the E4B Battlestaff Compartment, NCA Compartment, and Briefing Room are listed in Tables A4, A5, and A6. The overall quality score is a combination of two component scores, a score for signal quality (DAM/CSA) and a score for background quality (DAM/CBA). (See Section 3.3.) Tables of these component scores obtained for each of the acoustic environments are presented in Tables A7 through A12. The scores for overall quality were analyzed, and linear regression models for the relationship between quality scores and bit error rate were calculated. The quality scores obtained for individual speakers are shown as scatter diagrams in relation to the regression lines and the 95 percent confidence bounds for the data. For data for the Battlestaff Compartment, see Figures 10 and 11; for the NCA Compartment, see Figures 12 and 13; and for the Briefing Room data, see Figure 14. A comparison of the regression lines reveals the similarity of the three acoustic environments in their effects on vocoder quality. This similarity is illustrated in Figure 15.

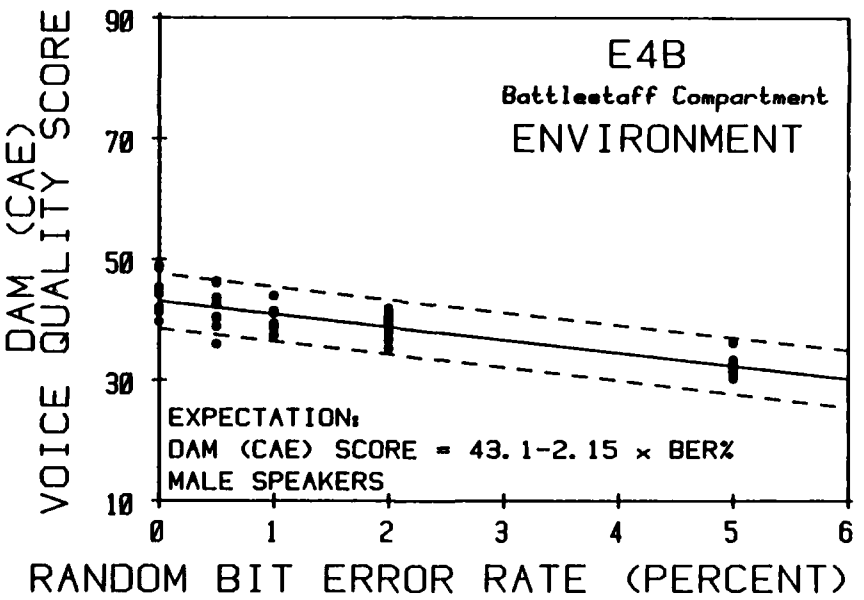


Figure 10. Male Speaker's Speech Quality (DAM/CAE) Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B Battlestaff Compartment Acoustic Environment

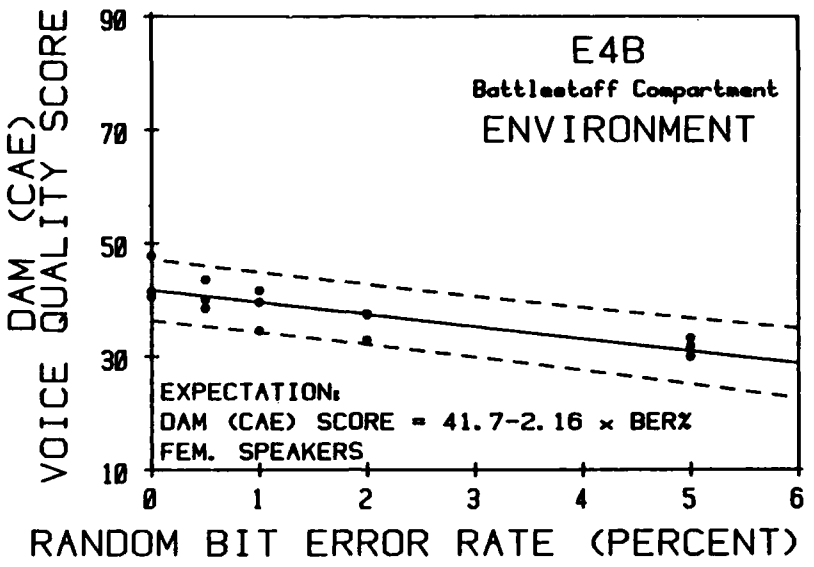


Figure 11. Female Speaker's Speech Quality (DAM/CAE) Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B Battlestaff Compartment Acoustic Environment

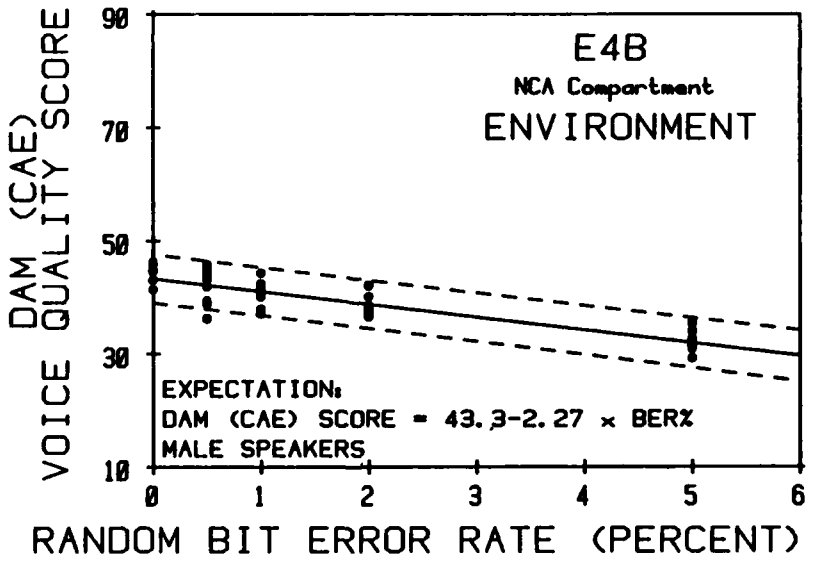


Figure 12. Male Speaker's Speech Quality (DAM/CAE) Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B NCA Compartment Acoustic Environment

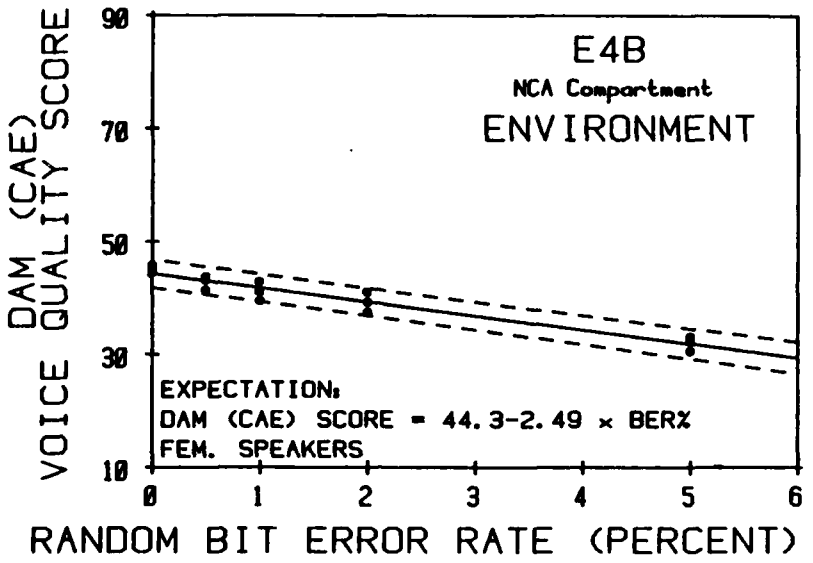


Figure 13. Female Speaker's Speech Quality (DAM/CAE) Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B NCA Compartment Acoustic Environment

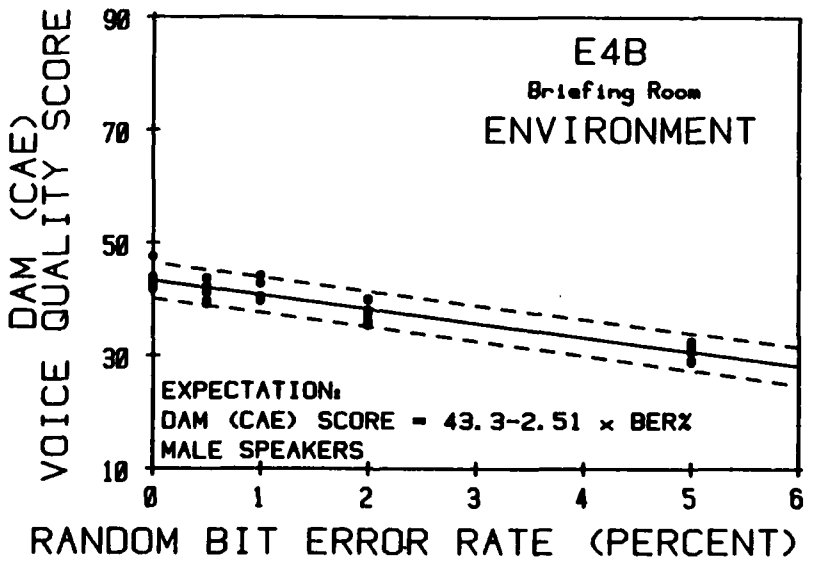


Figure 14. Male Speaker's Speech Quality (DAM/CAE) Scores vs Bit Error Rate, in Relation to Linear Regression Modeling, for the DOD LPC-10 Vocoder in the E4B Briefing Room Acoustic Environment

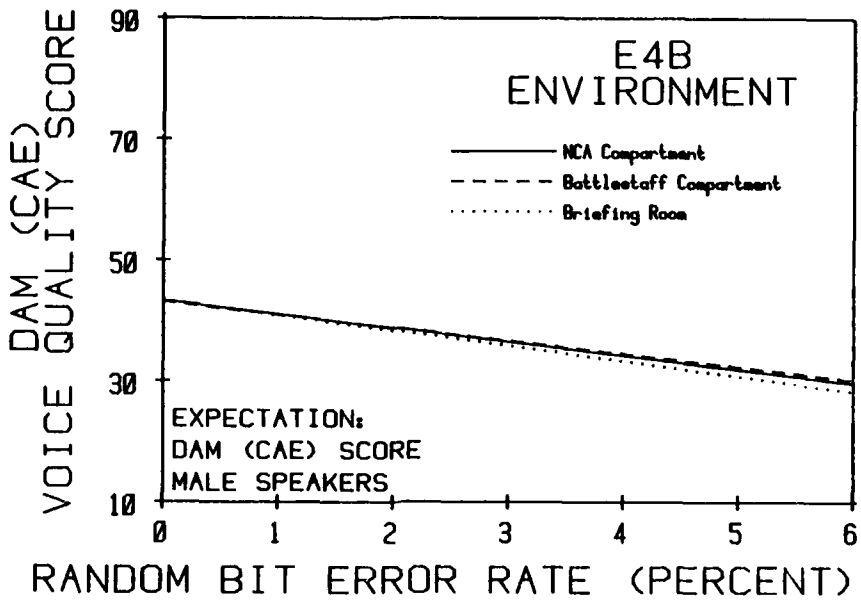


Figure 15. Comparison of Expected Variation in Speech Quality (DAM/CAE Scores) With Bit Error Rate for the DOD LPC-10 Vocoder Tested in the Acoustic Environments of the Three E4B Compartments

7.1 Composite Result: Voice Quality vs Bit Error Rate

Overall voice quality scores obtained for the vocoder in the three acoustic environments were pooled, and a regression model was calculated for the composite data. The resulting linear regression line and 95 percent confidence bounds for the data population are shown in Figure 16. With no bit errors, average quality score for the DOD LPC-10 vocoder in the composite acoustic environment is expected to be 43.3 (fair). Under bit error conditions, the average score is expected to fall 2.3 points for each percent bit error rate. Thus, at a 1-percent error rate, an average quality score of 41.0 (poor) is estimated.

7.2 Composite Result: Individual Speaker's Voice Quality Scores

Statistical tests comparing regression models fitted to voice quality scores vs error rate for individual speakers indicated that slopes of the regression lines calculated for individual speakers were not significantly different, but the elevations (origins) of regression lines for individual speaker's scores did differ by significant amounts. A new regression model and origins of the regression lines representing individual speaker's scores were calculated based on a common slope. This result is presented in Figure 17. The origins, estimates of the voice quality scores of

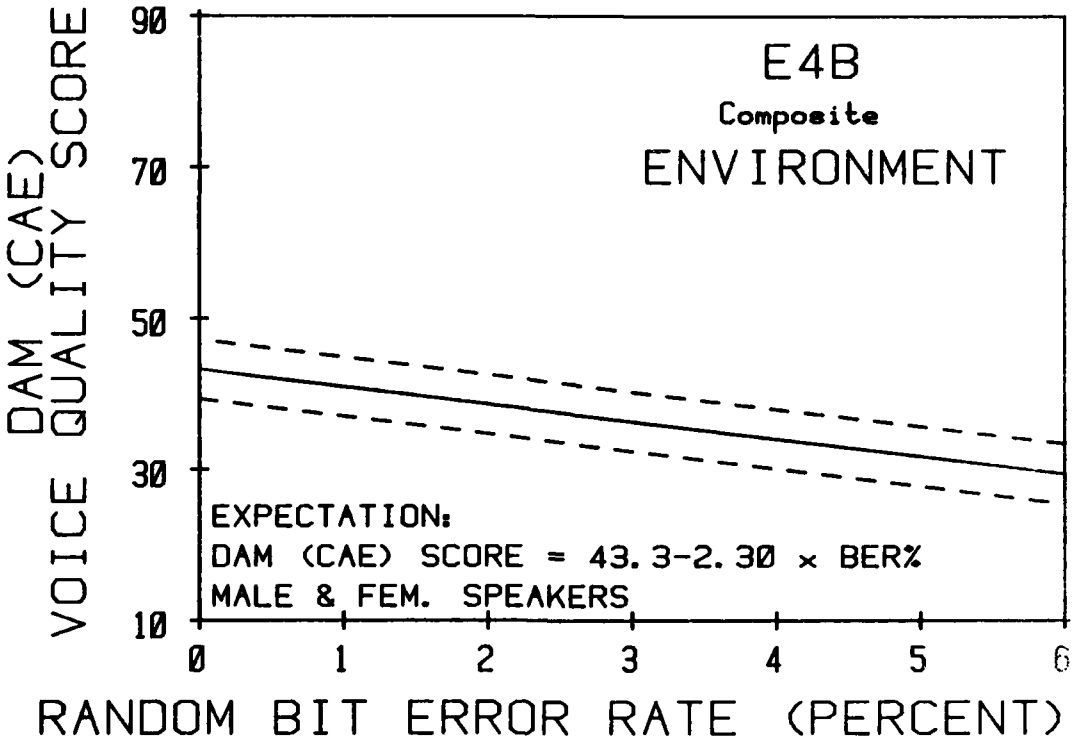


Figure 16. Linear Regression Model Expressing Variation in Speech Quality (DAM/CAE Scores) With Bit Error Rate, With 95% Confidence Limits for Individual Speaker's Voice Quality Scores, for the DOD LPC-10 Vocoder in the Composite E4B Acoustic Environment

individual speakers with no bit errors, varied from 45.4 for the best speaker to 41.3 for the poorest, with a common slope of -2.3.

8. CATEGORIES OF INTELLIGIBILITY AND VOICE QUALITY PERFORMANCE FOR THE DOD LPC-10 VOCODER IN THE E4B COMPOSITE ACOUSTIC ENVIRONMENT

The expected variation in intelligibility and voice quality of the DOD LPC-10 vocoder in the E4B acoustic environment is summarized in Table 5, which presents numerical values and categories for expected intelligibility performance (as shown in Figure 9) and voice quality performance (as shown in Figure 16). The statistical models fitted to the performance data indicate that, with no bit errors, an average intelligibility score of 80.2 can be expected, and that 95 percent of individual speaker's intelligibility scores will be distributed between 87.7 (good) and

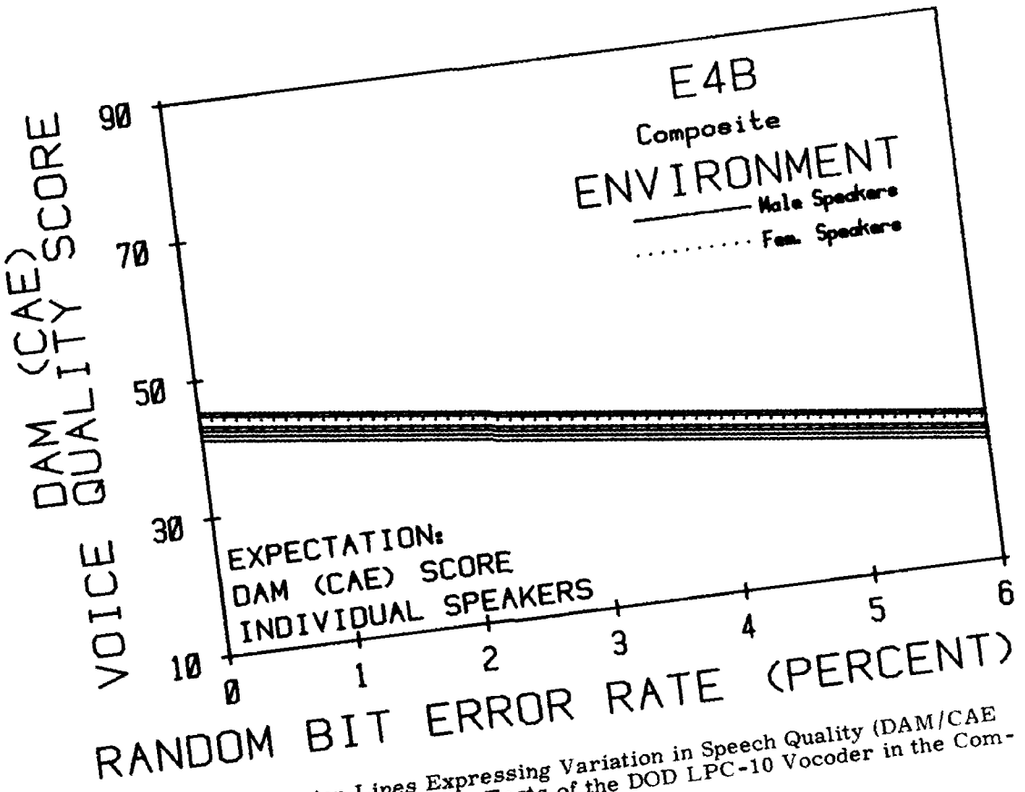


Figure 17. *Regression Lines* Expressing Variation in Speech Quality (DAM/CAE Scores) for Individual Speakers, in Tests of the DOD LPC-10 Vocoder in the Composite E4B Acoustic Environment

72.7 (very poor). The data indicate that 95 percent of individual speaker's intelligibility scores (assuming a population of which these speakers are a representative sample) will attain levels above the unacceptable threshold value of 70 when the bit error rate is below 0.7 percent. At that error rate, the average intelligibility is estimated to be 77 (poor). With no bit errors, an average voice quality score of 43.3 (fair) can be expected. The 95 percent confidence bounds for individual speaker's quality scores extend from 47.2 (fair) to 39.3 (poor). The data indicate that 95 percent of the individual speakers represented by this sample will attain quality levels above the unacceptable threshold value of 30 with bit error rates less than 4 percent. At the 4-percent error rate, the average level of quality is expected to be 34.1 (very poor), and the upper 95 percent confidence bound for the population is a score of 38.0 (poor).

Table 5. Predicted Scores and Categories of Performance for the DOD LPC-10 Vocoder in the E4B (Composite) Acoustic Environment, With the Resident Microphone With 95% Confidence Bounds for Scores From a Population of Male and Female Speakers

Bit Error Rate: 0	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%	3.5%	4.0%	4.5%	5.0%
INTELLIGIBILITY:										
Upper 95%	87.7	85.4	83.2	80.9	78.7	76.4	74.1	71.9	69.7	67.4	65.2
Bound	(good)-/--(moderate)----/-(fair) / (poor) ---------- (very poor)-/- UNACCEPTABLE--------										
Est. Avg Score	80.2	77.9	75.7	73.4	71.2	68.9	66.6	64.4	62.1	59.9	57.6
	(fair)-/--(poor)----------(very poor)--/-UNACCEPTABLE----------										
Lower 95%	72.7	70.4	68.2	65.9	63.7	61.4	59.1	56.9	54.6	52.3	50.1
Bound	(very poor)------/-UNACCEPTABLE----------										
OVERALL VOICE QUALITY:										
Upper 95%	47.2	46.0	44.9	43.7	42.6	41.4	40.3	39.1	38.0	36.8	35.7
Bound	(fair)---------- /-(poor) ----------/-(very poor)										
Est. Avg Score	43.3	42.1	41.0	39.8	38.7	37.5	36.4	35.2	34.1	32.9	31.8
	(fair)-------- / -(poor)----------/-(very poor)----------										
Lower 95%	39.3	38.2	37.0	35.9	34.7	33.6	32.4	31.3	30.1	29.0	27.8
Bound	(poor)---------- /-(very poor)----------/-UNACCEPTABLE										

9. COMPARATIVE PERFORMANCE: THE DOD LPC-10 NARROWBAND VOCODER UNDER QUIET CONDITIONS

It is instructive to compare the measured performance of the DOD standard LPC-10 vocoder under "ideal" conditions, that is, with speakers using a high-quality dynamic microphone in a quiet environment, with the measurements of performance obtained in the E4B acoustic environment. DRT intelligibility scores from earlier tests of this vocoder under quiet conditions are listed in Table A13, and voice quality ratings (DAM/CAE scores) are listed in Table A14. Separate components of the overall quality score, the scores for signal quality (DAM/CSA), and the scores for background quality (DAM/CBA) are listed in Tables A15 and A16.

9.1 Vocoder Intelligibility Under Ideal Conditions

A linear regression model calculated for the relationship between intelligibility and bit error rate from pooled intelligibility scores for male and female

speakers resulted in the regression line and 95 percent confidence bounds for the data shown in Figure 18. An average intelligibility score of 86.8 (moderate) is

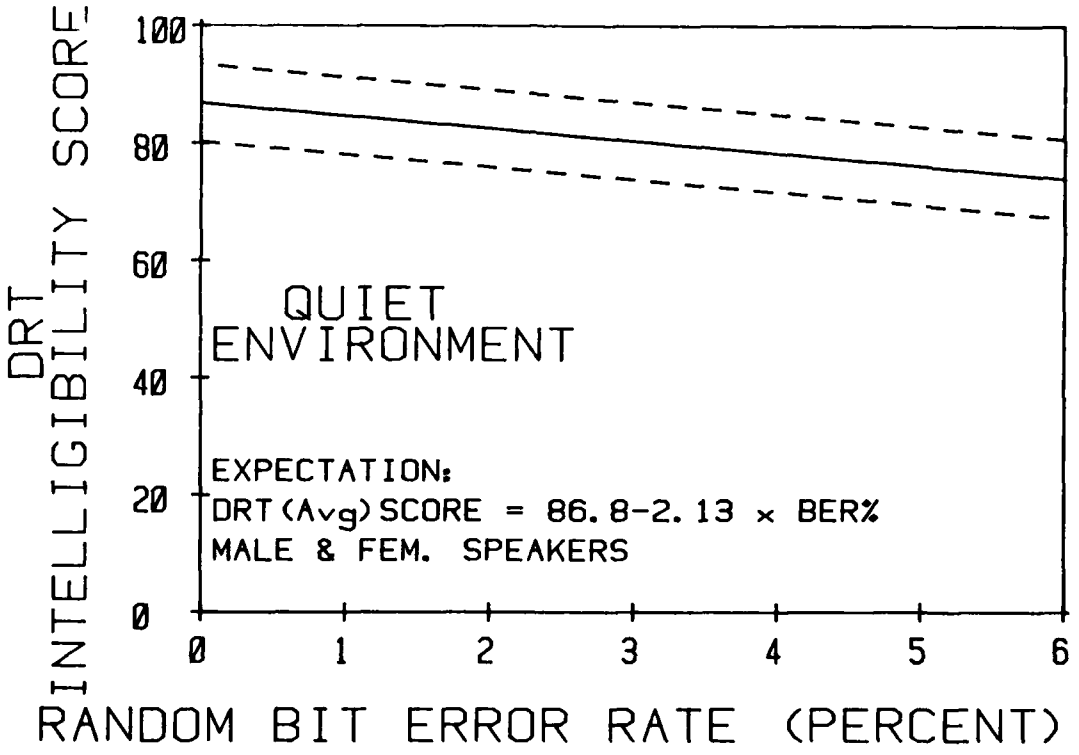


Figure 18. Linear Regression Model Expressing Variation in Intelligibility With Bit Error Rate, With 95% Confidence Limits for Individual Speaker's Intelligibility Scores, for the DOD LPC-10 Vocoder in a Quiet Environment

expected with no bit errors, with a regression line having a slope of -2.1. Statistical tests comparing the regression models calculated for individual speaker's scores in relation to bit error rate indicated that slopes of the individual regression functions did not differ significantly. However, the origins of the lines for individual speaker's scores differed by significant amounts. A new calculation based on a common slope of -2.13 established values for the origins of the regression lines for the six male and six female speakers' scores. The origins, representing the expected intelligibility with no bit errors, varied from 91.6 for the best speaker to 82.2 for the poorest. The average intelligibility of male speakers was almost 5 points higher than that of female speakers. This is illustrated in Figure 19.

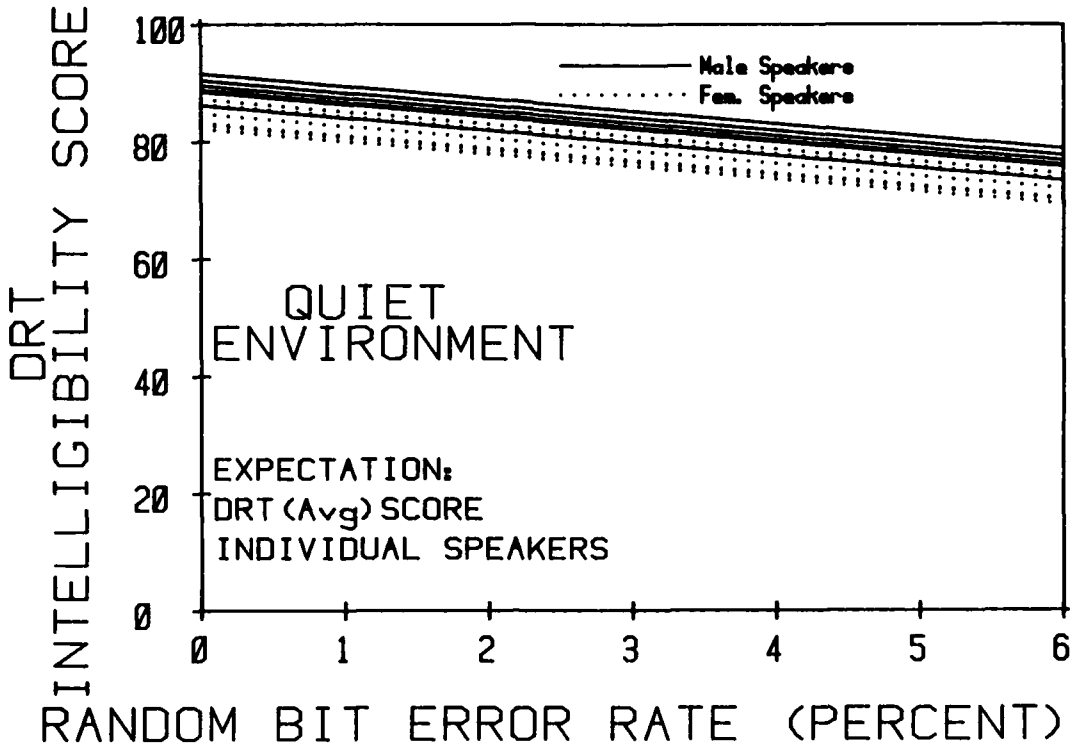


Figure 19. Regression Lines Expressing Variation in Intelligibility With Bit Error Rate of Individual Speakers, in Tests of the DOD LPC-10 Vocoder in a Quiet Environment

9.2 Vocoder Voice Quality Ratings (DAM/CAE Scores) Under Ideal Conditions

For the pooled data representing quality of the DOD LPC-10 vocoder tested with male and female speakers under quiet conditions, a linear regression model was calculated for the relationship between voice quality scores and bit error rate. The regression line and 95 percent confidence bounds for the data are shown in Figure 20. With no bit errors, the expected average quality score is 46.0 (fair), with a regression line slope of -2.6. The regression lines expressing variation of quality scores with bit error rate of individual speakers were found to have a common slope but significantly different origins, and a new regression model based on this finding established the variation of quality with bit error rate for individual speakers. These regression lines are shown in Figure 21. The origins, estimating voice quality scores with no bit errors, varied from 47.7 for the best speaker to 44.9 for the poorest, with a common slope of -2.62. The average quality scores for male speakers were 2.1 points higher than those for female speakers.

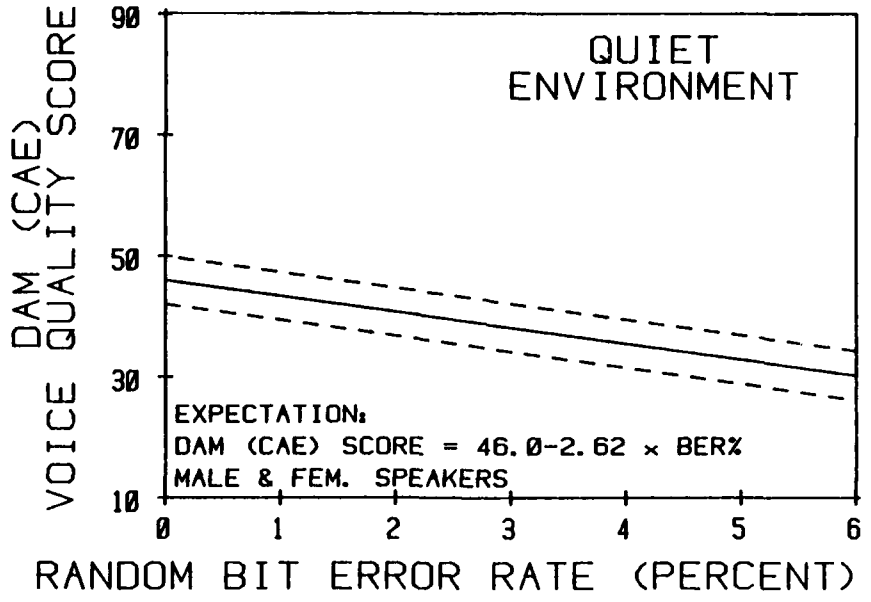


Figure 20. Linear Regression Model Expressing Variation in Speech Quality (DAM/CAE Scores), With 95% Confidence Limits for Individual Speaker's Voice Quality Scores, for the DOD LPC-10 Vocoder in a Quiet Environment

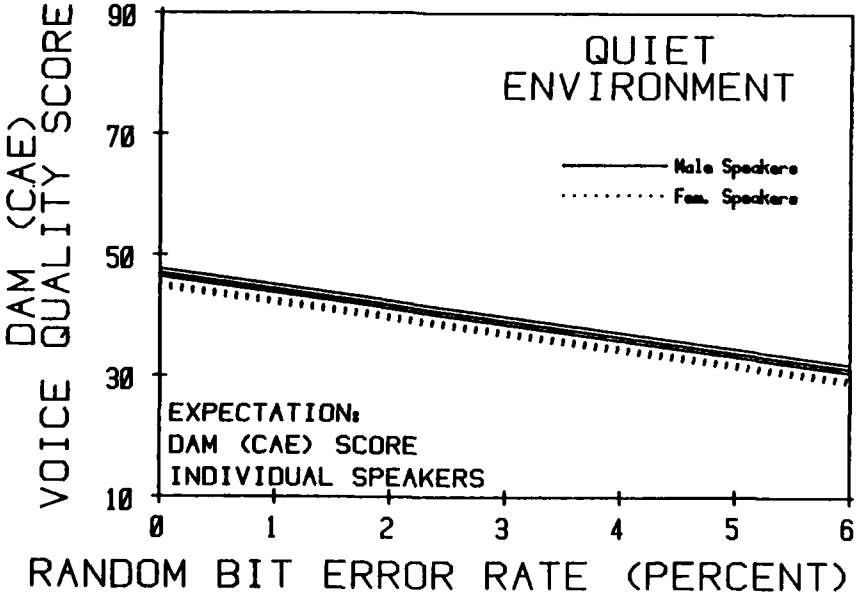


Figure 21. Regression Lines Expressing Variation in Speech Quality (DAM/CAE Scores) With Bit Error Rate, for Individual Speakers, in Tests of the DOD LPC-10 Vocoder in a Quiet Environment

9.3 Categories of Intelligibility and Voice Quality Performance for the DOD LPC-10 Vocoder in a Quiet Environment

The linear regression statistical models expressing variation in performance with bit error rate revealed in the data led to the results summarized in Table 6, which lists numerical values that correspond to the regression lines and confidence bounds presented in Figures 18 and 20. With no bit errors, it is expected that individual speaker's intelligibility scores for the DOD LPC-10 vocoder in a quiet environment will have 95 percent probability of being between 93.4 (very good) and 80.2 (fair). It is expected that 95 percent of the individual speaker's intelligibility scores from a population of which this was representative will be distributed above the unacceptable threshold value of 70 for error rates less than 4-1/2 percent. At the 4-1/2 percent error rate, an average intelligibility score of 77.2 (poor) is forecast, and the upper 95 percent boundary for individual's scores is 83.9 (moderate). The statistical model fitted to the voice quality data indicated that, with no bit errors, 95 percent of the scores would be distributed between 49.9 (moderate) and 42.0 (poor). At a 4-1/2 percent error rate, it is expected that average quality

Table 6. Predicted Scores and Categories of Performance for the DOD LPC-10 Vocoder in a Quiet Acoustic Environment (Dynamic Microphone) With 95% Confidence Bounds for Scores From a Population of Male and Female Speakers

Bit Error Rate: 0	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%	3.5%	4.0%	4/5%	5.0%
INTELLIGIBILITY:										
Upper 95% Bound	93.4	92.3	91.3	90.2	89.1	88.1	87.0	85.9	84.9	83.9	82.8
	(very good)----------/- (good) ----------/- (moderate)----------/- (fair)										
Est. Avg Score	86.8	85.8	84.7	83.6	82.6	81.5	80.4	79.4	78.3	77.2	76.2
	(moderate)---------- /-(fair)----------/-(poor)----------										
Lower 95% Bound	80.2	79.2	78.1	77.1	76.0	74.9	73.8	72.8	71.7	70.6	69.5
	(fair)---------/- (poor) ---------- /- (very poor)----------/-UNACCEPTABLE										
OVERALL VOICE QUALITY:										
Upper 95% Bound	49.9	48.6	47.3	46.0	44.7	43.3	42.0	40.7	39.4	38.2	36.9
	(moderate) ----- /-(fair) ----------/-(poor)----------										
Est. Avg Score	46.0	44.6	43.3	42.0	40.7	39.4	38.1	36.8	35.5	34.2	32.8
	(fair)---------- /-(poor)---------- /-(very poor)----------										
Lower 95% Bound	42.0	40.7	39.4	38.1	36.8	35.4	34.1	32.8	31.5	30.2	28.8
	(poor)---------- /-(very poor) ---------- /-UNACCEPTABLE										

will be 34.2 (very poor), with 95 percent of the scores of individual speakers represented by this sample distributed between 38.2 (poor) and 30.2 (very poor).

10. COMPARATIVE PERFORMANCE: WIDEBAND CONTINUOUS VARIABLE-SLOPE DELTA MODULATION (CVSD) VOICE PROCESSORS

Previous studies have established intelligibility scores and voice quality ratings of wideband CVSD processors operating at 16K bps and 32K bps data rates. Those studies were for ideal conditions, that is, speakers were in a quiet environment.

10.1 *Intelligibility of the CVSD-16 Processor in a Quiet Environment*

DRT intelligibility scores obtained from tests of a 16K bps CVSD processor at 0, 1/2, 1, 2, and 5 percent bit error rates are listed in Table A16. These data were used in calculating a linear regression model leading to the regression line and 95 percent confidence bounds for the data shown in Figure 22. With no bit errors, the average intelligibility is expected to be 91.4 (very good). The calculated

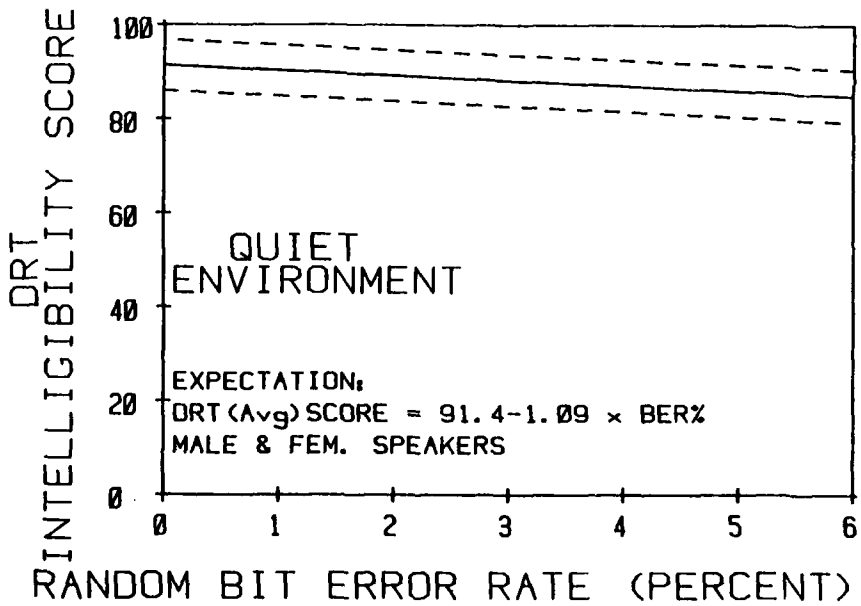


Figure 22. Linear Regression Model Expressing Variation in Intelligibility With Bit Error Rate, With 95% Confidence Limits for Individual Speaker's Intelligibility Scores, for the Wideband CVSD-16 Processor in Quiet Environment

slope of the regression function is -1.1. Regression functions modeling individual speaker's intelligibility vs bit error rate were found to have a common slope, but differ significantly in the origins of the lines. A regression model calculated on the basis of a common slope determined regression functions for individual speakers, illustrated in Figure 23. The origins, representing the expected intelligibility score with no bit errors, ranged from 94.9 for the best speaker to 89.0 for the poorest. Male speakers' intelligibility scores tended to exceed those of female speakers.

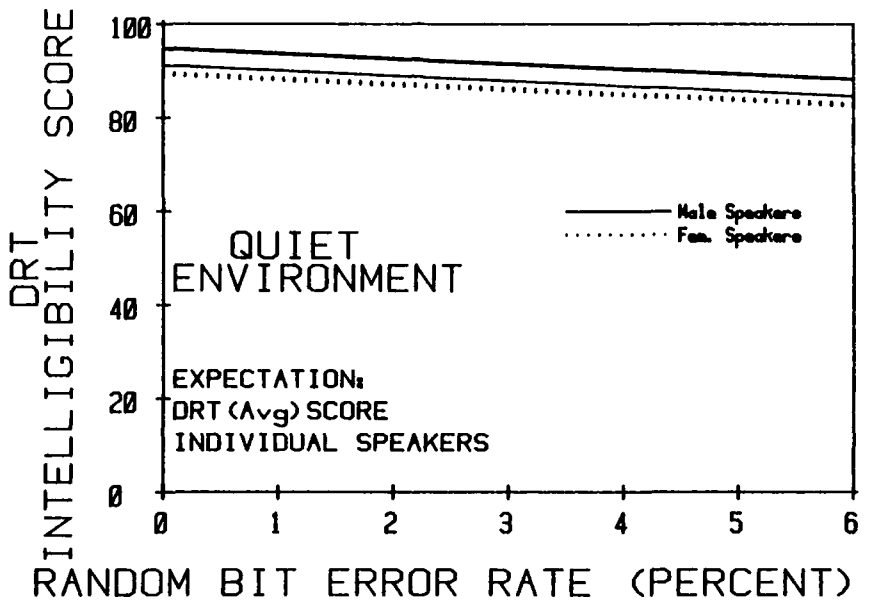


Figure 23. Regression Lines Expressing Variation in Intelligibility With Bit Error Rate of Individual Speakers, in Tests of the Wideband CVSD-16 Processor in Quiet Environment

10.2 Voice Quality Ratings (DAM/CAE Scores) for the CVSD-16 Processor in a Quiet Environment

Voice quality scores from three independent tests of the CVSD-16 processor are listed in Table A17. Corresponding scores for the components of the overall quality score, the scores for signal quality (DAM/CSA), and the scores for background quality (CAM/CBA) are listed in Tables A18 and A19. The overall quality (DAM/CAE) scores for male and female speakers were pooled, and a regression model was calculated for the combined data, leading to the result illustrated in Figure 24. The figure presents the regression line and 95 percent confidence

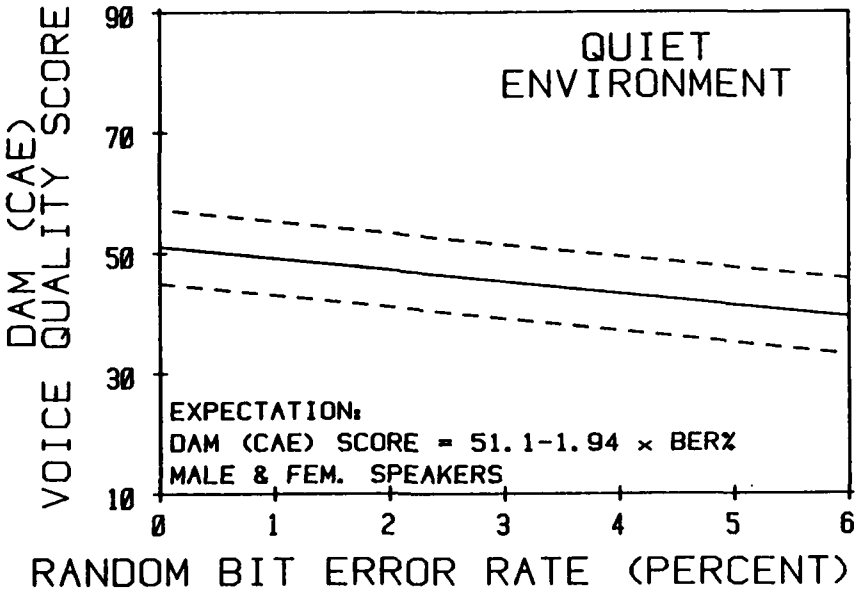


Figure 24. Linear Regression Model Expressing Variation in Speech Quality (DAM/CAE Scores) With Bit Error Rate, With 95% Confidence Limits for Individual Speaker's Voice Quality Scores, for the Wideband CVSD-16 Processor in a Quiet Environment

bounds for this collection of scores. The statistical model indicated that, with no bit errors, an average quality score of 51.1 (moderate) was obtained, with a slope of -1.9 for the regression line. Regression lines calculated for scores vs bit error rate of individual speakers did not differ significantly in the slopes, but significant differences were indicated for the origins. A new regression model was calculated on the basis of a common slope. The resulting regression functions for individual speaker's voice quality scores vs error rate are shown in Figure 25. Origins of the lines, estimating individual speaker's voice quality scores with no bit errors, varied from 53.6 for the best speaker to 49.4 for the poorest, with a common slope of -1.94. Quality ratings for male speakers tended to exceed those obtained for female speakers.

10.3 Categories of Intelligibility and Voice Quality Performance for the CVSD-16 Processor

The statistical models fitted to the intelligibility data and voice quality data for the CVSD-16 processor, illustrated graphically in Figures 22 and 24, provided the numerical values presented in Table 7. The table gives the descriptive categories associated with the levels of performance. With no bit errors, it is estimated that

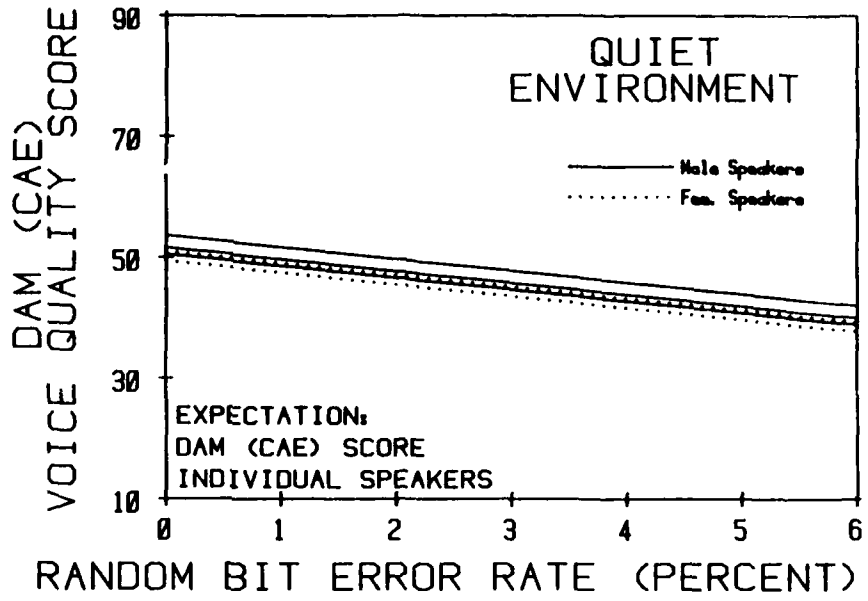


Figure 25. Regression Lines Expressing Variation in Speech Quality (DAM/CAE Scores) With Bit Error Rate of Individual Speakers, in Tests of the Wideband CVSD-16 Processor in a Quiet Environment

95 percent of the intelligibility scores of individual speakers from a population represented by this sample will be between 96.9 (excellent) and 86.0 (moderate). At a 3 percent error rate, the expected average intelligibility score is 88.2 (good), with an estimate that 95 percent of individual speaker's intelligibility scores will fall between 93.6 (very good) and 82.7 (fair). Voice quality performance fell into categories somewhat below those for intelligibility. With no bit errors, it is estimated that 95 percent of the quality scores of individual speakers will fall between 57.2 (good) and 44.9 (fair). At a 3 percent error rate, the expected average quality score is 45.2 (fair), and the 95 percent confidence bounds for individual speaker's quality scores extend from 51.4 (moderate) to 39.1 (poor).

10.1 Intelligibility of the CVSD-32 in a Quiet Environment

Intelligibility scores obtained from tests of the 32 K bps CVSD processor under bit error conditions in a quiet acoustic environment are listed in Table A20. A linear regression model calculated from these scores established the relationship between performance and bit error rate; the regression line and 95 percent confidence bounds for the data are shown in Figure 26. Without bit errors, an average intelligibility score of 93.6 (very good) is expected, with a regression slope of

Table 7. Predicted Scores and Categories of Performance for the CVSD-16 Processor in a Quiet Acoustic Environment (Dynamic Microphone) With 95% Confidence Bounds for Scores From a Population of Male and Female Speakers

Bit Error Rate: 0	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%	3.5%	4.0%	4.5%	5.0%
INTELLIGIBILITY:										
Upper 95% Bound	96.6	96.3	95.7	95.2	94.6	94.1	93.6	93.0	92.5	92.0	91.5
	(excellent)----/-(very good)----------										
Est. Avg. Score	91.4	90.9	90.3	89.8	89.2	88.7	88.2	87.6	87.1	86.5	86.0
	(very good)/(good)----------									/-(moderate)------	
Lower 95% Bound	86.0	85.5	84.9	84.4	83.8	83.3	82.7	82.2	81.6	81.0	80.5
	(moderate)----------						/-(fair)----------				
OVERALL VOICE QUALITY:										
Upper 95% Bound	57.2	56.2	55.2	54.3	53.3	52.3	51.4	50.4	49.5	48.5	47.6
	(good)----------				/-(moderate) ----------					/-(fair)	
Est. Avg. Score	51.1	50.1	49.1	48.1	47.2	46.2	45.2	44.3	43.3	42.3	41.3
	(moderate)----------			/-(fair) ----------						/-(poor)	
Lower 95% Bound	44.9	44.0	43.0	42.0	41.1	40.1	39.1	38.1	37.1	36.1	35.1
	(fair)----------		/-(poor) ----------							/-(very poor)	

-0.75; that is, scores are expected to drop by three-fourths of a point with each percent of bit errors. As with the other processors, the regression functions calculated for individual speaker's intelligibility scores vs bit error rate were found to have a common slope, but differ significantly in regard to the elevations of the regression lines. A new regression model was calculated on the basis of this finding. The result, shown in Figure 27, gave values of the origins of the regression lines that varied from 96.6 for the best speaker to 90.7 for the poorest.

10.5 Voice Quality Ratings (DAM/CAE Scores) for the CVSD-32 Processor in a Quiet Environment

Two independent tests of the voice quality of the CVSD-32 processor with input speech from a quiet acoustic environment resulted in the voice quality scores listed in Table A21. The regression model fitted to this data indicated the relationship between quality scores and bit error rate that is shown in Figure 28. With this group of speakers, an average quality score of 58.3 (very good) is expected with no bit errors, with a slope of the regression function of -2.4. Comparisons of re-

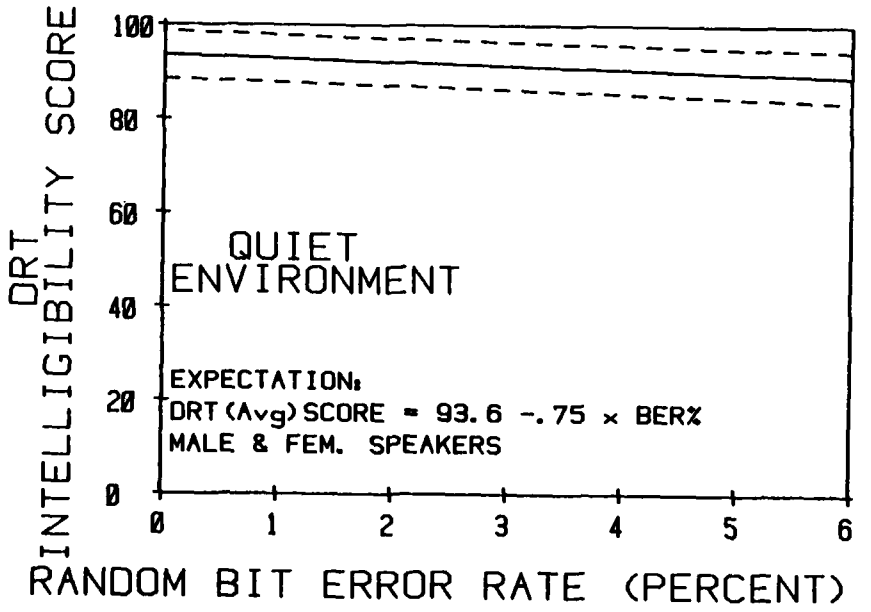


Figure 26. Linear Regression Model Expressing Variation in Intelligibility With Bit Error Rate, With 95% Confidence Limits for Individual Speaker's Intelligibility Scores, for the Wideband CVSD-32 Processor in a Quiet Environment

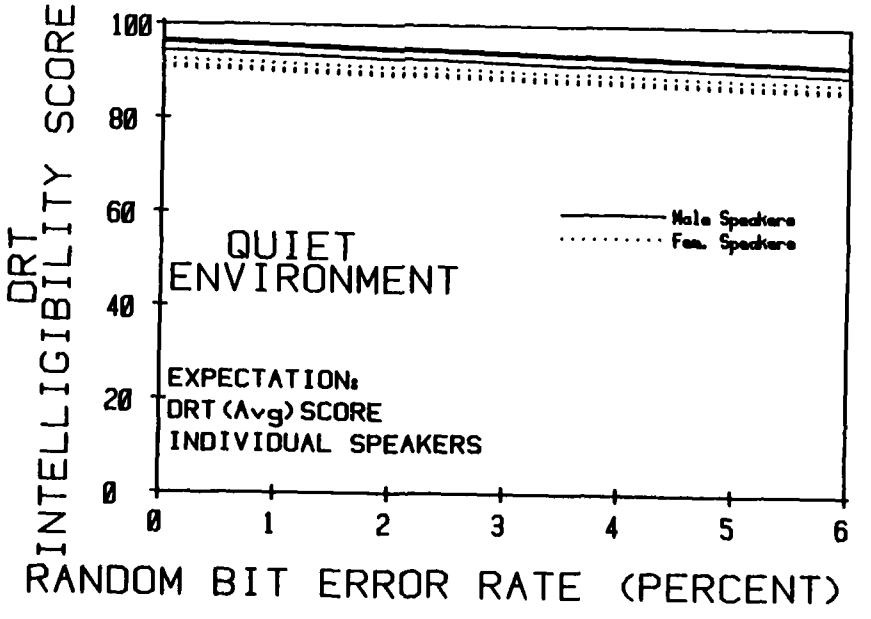


Figure 27. Regression Lines Expressing Variation in Intelligibility With Bit Error Rate for Individual Speaker's Intelligibility Scores for the Wideband CVSD-32 Processor in a Quiet Environment

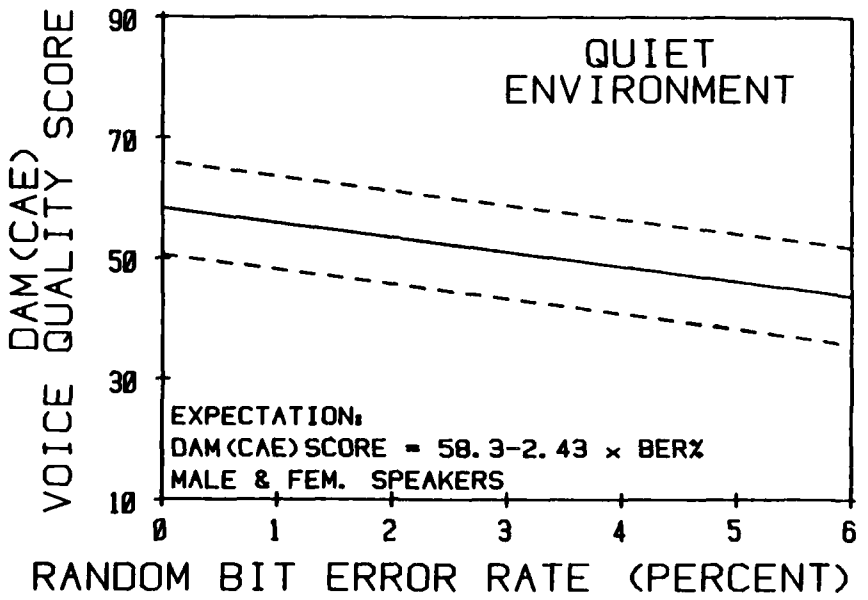


Figure 28. Linear Regression Model Expressing Variation in Speech Quality (DAM/CAE Scores) With Bit Error Rate, With 95% Confidence Limits for Individual Speaker's Voice Quality Scores, From Tests of the Wideband CVSD-32 Processor in a Quiet Environment

gression models calculated from individual speaker's scores indicated that, while the elevations of the regression lines based on individual speaker's scores differed significantly, the differences in slopes of the lines were not significant. A new regression model was calculated based on this finding, and the result is shown in Figure 29. Origins of regression lines for individual speakers, estimating the voice quality score with no bit errors, varied from 61.7 for the best speaker to 54.0 for the poorest, with a common slope of -2.43.

10.6 Categories of Intelligibility and Voice Quality Performance for the CVSD-32 Processor in a Quiet Environment

The statistical models for the variation in performance with bit error rate illustrated in Figures 26 and 28 provided the numerical values and associated descriptive categories shown in Table 8. With no bit errors, it is estimated that 95 percent of the intelligibility scores of individual speakers will fall between 98.7 (excellent) and 88.5 (good), and voice quality scores will fall between 66.1 (excellent) and 50.5 (moderate). At a 3 percent error rate, it is predicted that intelligibility scores will have 95 percent probability of falling between 96.4 (excellent) and 86.3 (moderate), while quality scores will fall between 58.8 (very good) and 43.3 (fair).

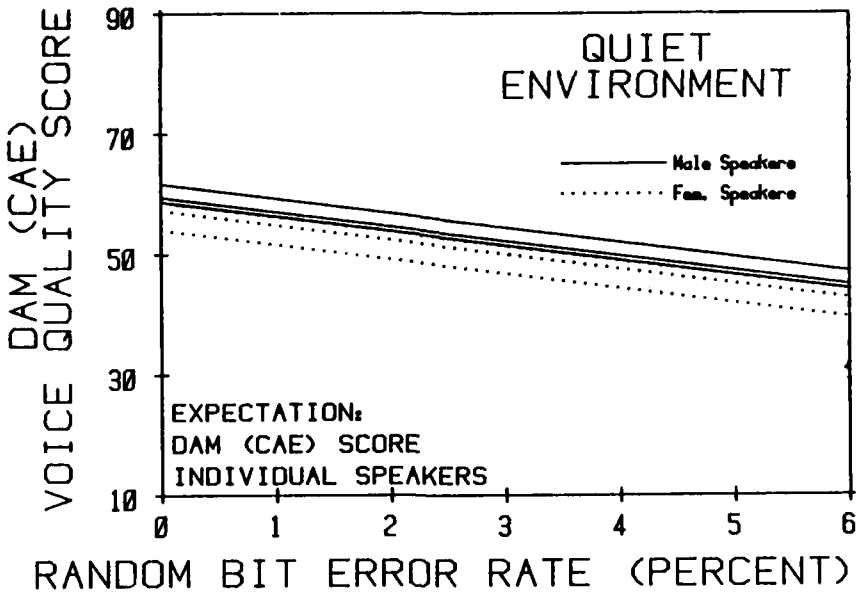


Figure 29. Regression Lines Expressing Variation in Speech Quality (DAM/CAE Scores) vs Bit Error Rate for Individual Speakers in Tests of the Wideband CVSD-32 Processor in Quiet Environment

11. COMPARISONS OF VOICE PROCESSORS AND TEST CONDITIONS

Comparisons of performance capabilities of the voice processors are facilitated with the linear regression models for performance vs bit error rate. In Figure 30, predicted intelligibility performance is compared with bit error rate for the DOD narrowband LPC-10 vocoder in the E4B acoustic environment and in a quiet environment. Results from tests of an LPC-10 vocoder lacking bit error detection and correction are also shown for comparison,[1] and for the functions estimating performance of the CVSD-16 and CVSD-32 processors. Comparisons of voice quality performance with bit error rate are presented in Figure 31. Relative ranking of the processors is the same according to intelligibility and voice quality. With increasing bit error rate, however, the differences in quality of the LPC-10 vocoder in the quiet and in the E4B environment tend to become smaller, while the differences in intelligibility performance tend to become larger. The linear regression coefficients are compared in Table 9, which also lists the coefficients calculated for the best and the poorest speakers for each processor.

Table 8. Predicted Scores and Categories of Performance for the CVSD-32 Processor in a Quiet Acoustic Environment (Dynamic Microphone) With 95% Confidence Bounds for Scores From a Population of Male and Female Speakers

Bit Error Rate: 0	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%	3.5%	4.0%	4.5%	5.0%
INTELLIGIBILITY:										
Upper 95% Bound	98.7	98.3	97.9	97.5	97.1	96.7	96.4	96.0	95.7	95.4	95.1
	(excellent) ---------- /-(very good) ----------										
Est. Avg Score	93.6	93.2	92.8	92.5	92.1	91.7	91.3	90.9	90.6	90.2	89.8
	(very good) ---------- /-(good) ----------										
Lower 95% Bound	88.5	88.2	87.8	87.4	87.1	86.7	86.3	85.9	85.4	85.0	84.5
	(good) ---------- /-(moderate) ----------										
OVERALL VOICE QUALITY:										
Upper 95% Bound	66.1	64.8	63.6	62.4	61.2	60.0	58.8	57.6	56.4	55.2	54.1
	(excellent) ----/-(very good) ----------/-(good) ---------- /-(moderate)										
Est. Avg Score	58.3	57.1	55.9	54.7	53.4	52.2	51.0	49.8	48.6	47.4	46.2
	(very good)/(good) ---------- /-(moderate) ----------/-(fair) ---------										
Lower 95% Bound	50.5	49.4	48.2	47.0	45.7	44.5	43.3	42.0	40.8	39.5	38.2
	(moderate) ----------/-(fair) ----------/-(poor) ----------										

11.1 Diagnostic Intelligibility Scores vs Bit Error Rate for the DOD LPC-10 Vocoder

The variations in scores obtained for the separate intelligibility features of the DRT provide insights regarding the relative vulnerability of the different phonetic features to bit error effects and noise effects. For this reason, separate linear regression models were calculated for each of the intelligibility features tested by the DRT. The results from the performance data on the vocoder in the E4B composite acoustic environment are summarized in Figure 32; the corresponding results for vocoder performance in a quiet environment are summarized in Figure 33. For both conditions, the feature "sibilation" obtained the highest scores, and the features "graveness" and "sustention" received the poorest scores, but were under much greater impairment in the noise environment. Details of the variations in scores for sustention are shown in Figure 34, and, for graveness, in Figure 35.

Research on improving the intelligibility of these phonetic features could provide the most payoff in overall improvement in intelligibility performance of the vocoder under these impairments. The regression coefficients determined for individual features are compared in Table 10.

Table 9. Comparison of Voice Processors: Average Performance vs Bit Error Rate for a Population of Male and Female Speakers

Processor, Environment, and Microphone		Intelligibility Estimate (Avg DRT Score)	Voice Quality Estimate (Avg DAM/CAE Score)
DOD LPC-10 Vocoder	Average speaker:	S(DRT) = 80.2-4.51xBER(%)	S(DAM/CAE) = 43.3-2.30xBER(%)
in E4B (Composite)	Best:	= 84.4-4.51xBER(%)	= 45.4-2.30xBER(%)
Environment, with	Poorest:	= 77.2-4.51xBER(%)	= 41.3-2.30xBER(%)
resident microphone			
DOD LPC-10 Vocoder	Average speaker:	S(DRT) = 86.8-2.13xBER(%)	S(DAM/CAE) = 48.0-2.62xBER(%)
in Quiet, with	Best:	= 91.6-2.13xBER(%)	= 47.7-2.62xBER(%)
dynamic microphone	Poorest:	= 82.2-2.13xBER(%)	= 44.9-2.62xBER(%)
CVSD-16 Processor	Average speaker:	S(DRT) = 91.4-1.09xBER(%)	S(DAM/CAE) = 51.1-1.94xBER(%)
in Quiet, with	Best:	= 94.9-1.09xBER(%)	= 53.6-1.94xBER(%)
dynamic microphone	Poorest:	= 89.0-1.09xBER(%)	= 49.4-1.94xBER(%)
CVSD-32 Processor	Average speaker:	S(DRT) = 93.6-0.75xBER(%)	S(DAM/CAE) = 58.3-2.43xBER(%)
in Quiet, with	Best:	= 96.6-0.75xBER(%)	= 61.7-2.43xBER(%)
dynamic microphone	Poorest:	= 90.7-0.75xBER(%)	= 54.0-2.43xBER(%)

11.2 Signal Quality (DAM/CSA) and Background Quality (DAM/CBA) Scores vs Bit Error Rate

The variations in signal quality (DAM/CSA) scores with bit error rate are compared for the LPC-10 vocoder in the E4B acoustic environment and the vocoder in quiet in Figure 36. The variation in background quality (DAM/CBA) scores with bit error rate are compared in Figure 37. Signal quality and background quality were both lowered in the E4B noise environment. The comparison of signal quality scores indicated higher scores for male speakers in the quiet, but this ranking was reversed in the E4B acoustic noise environment, where the female speakers tended to receive the higher scores for signal quality. The scores for background quality were virtually identical for male and female speakers.

12. DISCUSSION OF FINDINGS

The acoustic noise environment of the E4B caused a significant lowering of speech intelligibility scores and voice quality scores obtained for the DOD LPC-10 vocoder. Among the factors that have been found to affect intelligibility scores and voice quality scores are the particular microphones and the particular combinations of speakers used in conducting tests.

Table 10. Table of Linear Regression Coefficients Estimating Intelligibility vs Bit Error Rate for the DOD Narrowband LPC-10 Vocoder. A = origin of regression line, and B = slope

Intelligibility Feature	VOCODER IN QUIET (Dynamic Microphone)						VOCODER IN E4B (Composite) ENVIRONMENT (Resident Microphone)					
	MALES		FEMALES		MALES & FEMALES		MALES		FEMALES		MALES & FEMALES	
	A	B	A	B	A	B	A	B	A	B	A	B
VOICING												
Present	92.5	-1.61	89.4	-2.35	91.0	-1.98	81.9	-4.25	79.7	-4.41	81.4	-4.29
Absent	86.8	-1.70	91.0	-1.83	88.9	-1.80	83.3	-4.66	91.2	-2.77	85.3	-4.19
Pres. & Abs.	89.7	-1.66	90.2	-2.09	90.0	-1.88	82.6	-4.46	85.4	-3.59	83.3	-4.24
NASALITY												
Present	95.5	-2.34	98.2	-1.80	96.9	-2.07	83.1	-4.11	94.2	-4.42	85.9	-4.19
Absent	88.7	-3.93	78.1	-2.29	83.4	-3.11	82.5	-3.44	90.9	-2.28	84.6	-3.15
Pres. & Abs.	92.1	-3.13	77.2	-2.27	90.2	-2.59	82.8	-3.78	92.6	-3.35	85.2	-3.67
SUSTENTION												
Present	86.4	-1.69	81.8	-1.88	84.1	-1.70	63.6	-6.35	72.7	-7.98	65.8	-6.75
Absent	77.5	-3.58	72.6	-2.66	75.1	-3.12	62.1	-5.04	69.5	-4.49	64.0	-4.90
Pres. & Abs.	82.0	-2.64	77.2	-2.27	79.6	-2.46	62.8	-5.69	71.1	-6.24	65.0	-5.83
SIBILATION												
Present	95.7	-1.03	84.7	-0.78	90.2	-0.91	90.5	-5.97	76.3	-7.00	87.0	-6.23
Absent	96.9	-0.33	86.8	-2.05	91.9	-1.19	94.4	-3.58	87.2	-3.57	92.6	-3.58
Pres. & Abs.	96.3	-0.69	85.8	-1.41	91.1	-1.05	92.4	-4.77	81.8	-5.29	89.8	-4.91
GRAVENESS												
Present	86.7	-4.84	78.6	-3.83	82.7	-4.34	73.7	-5.50	71.8	-4.71	73.2	-5.31
Absent	76.9	-1.95	66.2	-1.02	71.6	-1.49	61.8	-3.46	63.5	-3.15	62.2	-3.38
Pres. & Abs.	81.8	-3.40	72.4	-2.43	77.1	-2.92	67.7	-4.48	67.7	-3.93	67.7	-4.35
COMPACTNESS												
Present	93.9	-1.05	94.2	-1.59	94.1	-1.32	90.8	-4.48	92.9	-2.73	91.3	-4.04
Absent	93.2	-2.03	91.2	-3.16	92.2	-2.60	88.7	-4.11	90.2	-4.35	89.1	-4.04
Pres. & Abs.	93.5	-1.54	92.7	-2.38	93.1	-1.96	89.7	-4.30	91.8	-3.54	90.2	-4.11

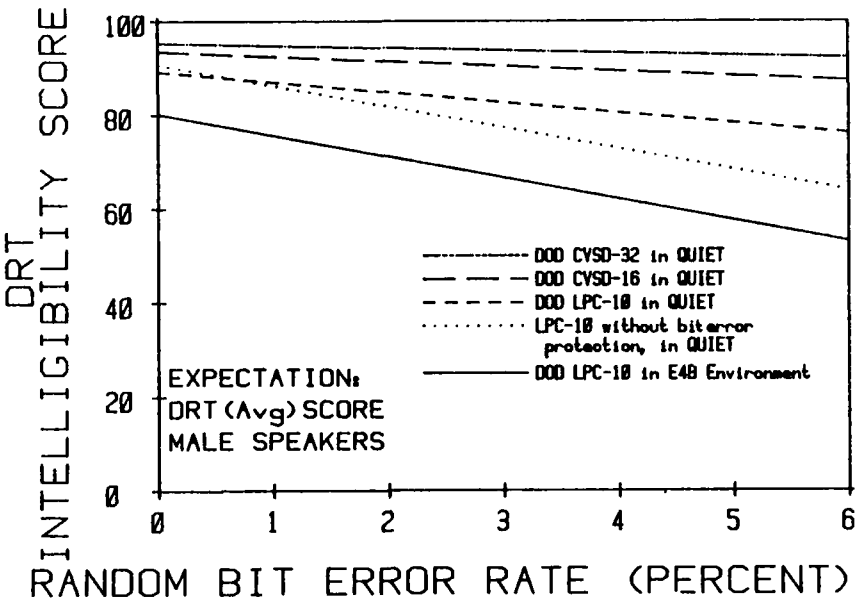


Figure 30. Comparison of Regression Lines Estimating Intelligibility vs Bit Error Rates for the DOD LPC-10 Vocoder in the E4B and Quiet Acoustic Environments With the CVSD-16 and CVSD-32 Processors

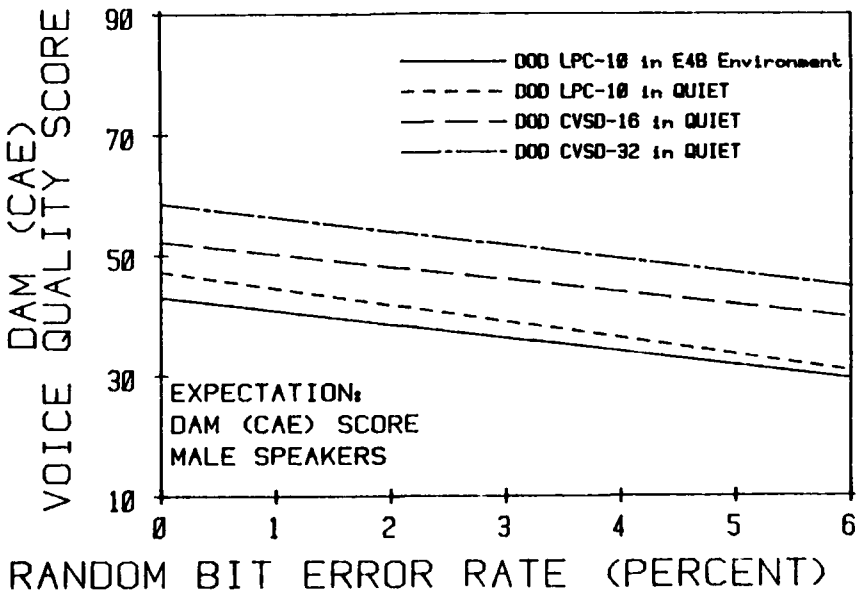


Figure 31. Comparison of Regression Lines Estimating Voice Quality vs Bit Error Rate for the DOD LPC-10 Vocoder in the E4B and Quiet Acoustic Environments, and the CVSD-16 and CVSD-32 Wideband Processors

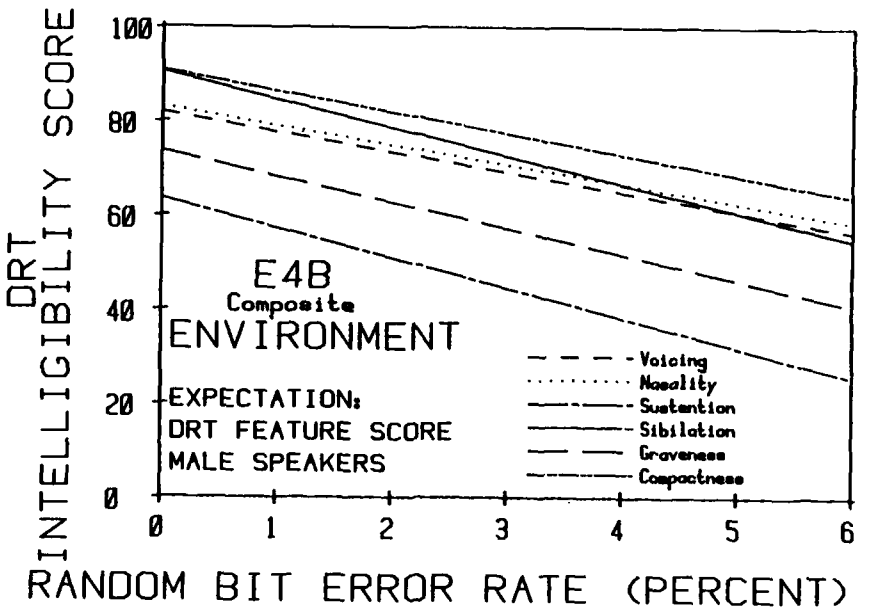


Figure 32. Comparison of Regression Lines Estimating Scores for the Intelligibility Features vs Bit Error Rate for the DOD LPC-10 Vocoder in the E4B Composite Acoustic Environment

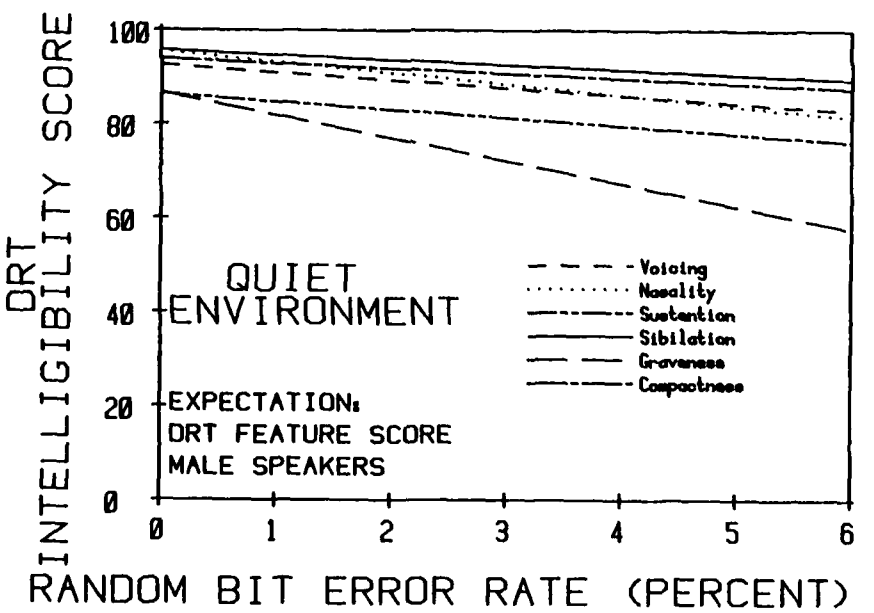


Figure 33. Comparison of Regression Lines Estimating Scores for the Intelligibility Features vs Bit Error Rate for the DOD LPC-10 Vocoder in a Quiet Acoustic Environment

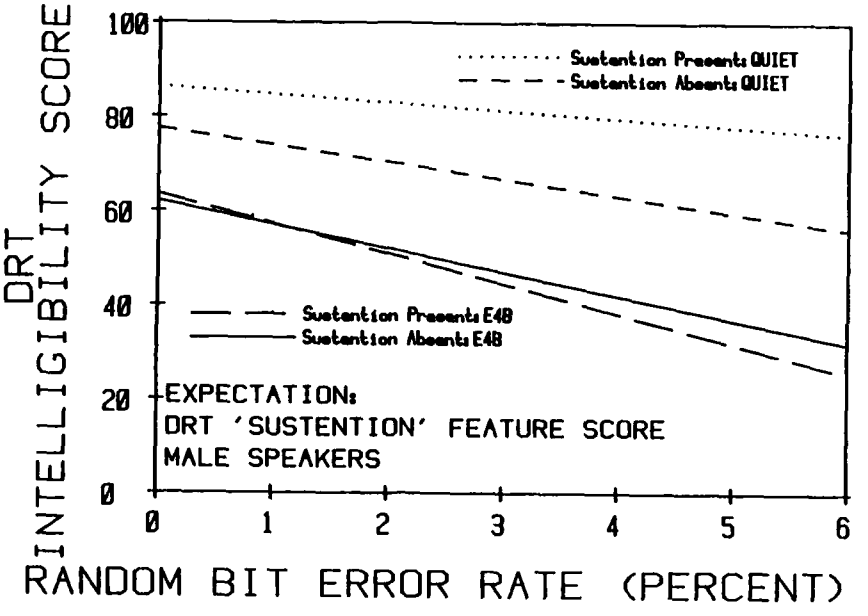


Figure 34. Comparison of Regression Lines Estimating Scores for the Sustention Intelligibility Feature vs Bit Error Rate for the DOD LPC-10 Vocoder in Quiet and E4B Environments

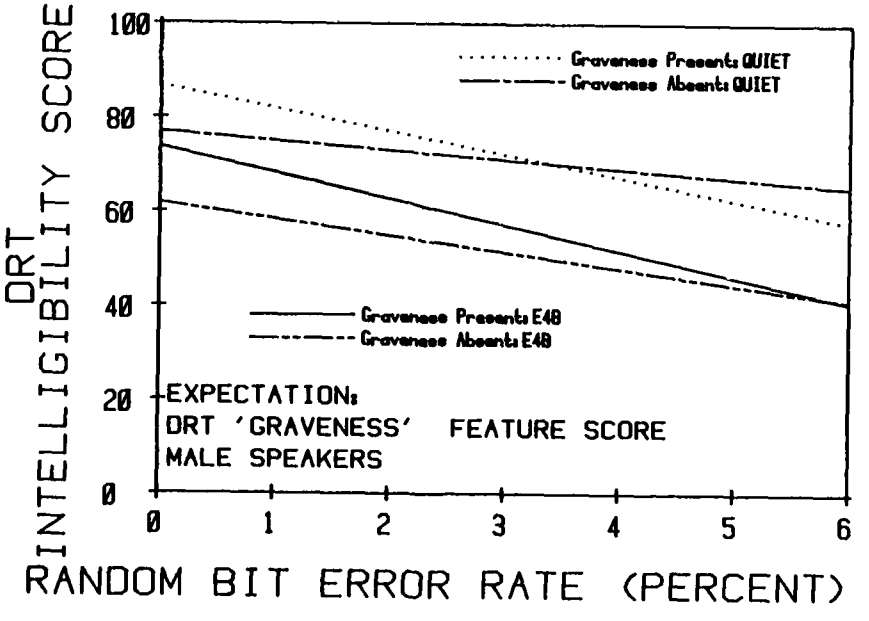


Figure 35. Comparison of Regression Lines Estimating Scores for the Graveness Intelligibility Feature vs Bit Error Rate for the DOD LPC-10 Vocoder in Quiet and E4B Environments

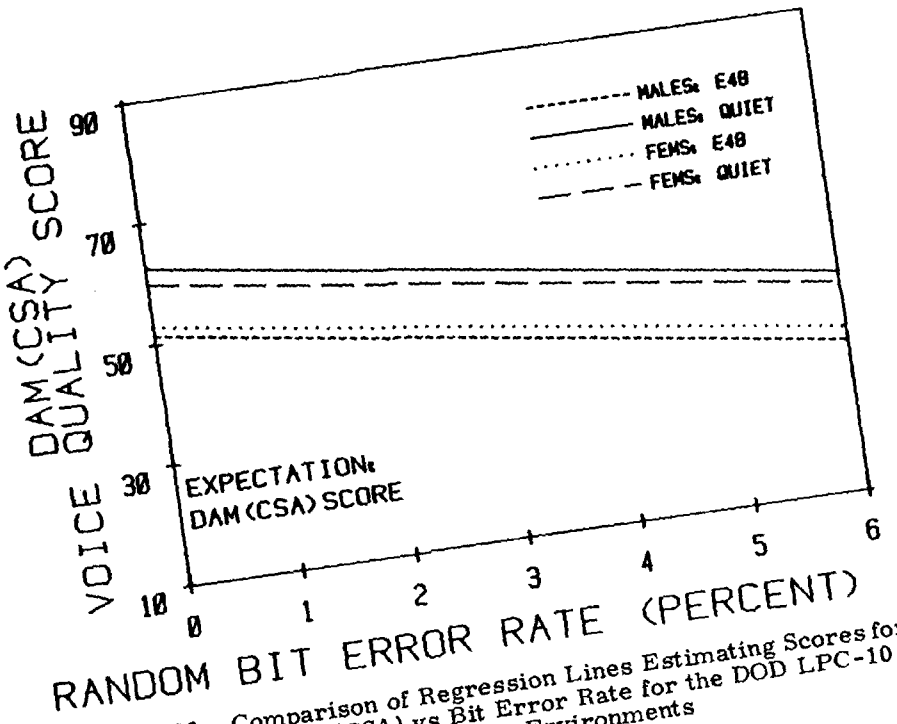


Figure 36. Comparison of Regression Lines Estimating Scores for Signal Quality (DAM/CSA) vs Bit Error Rate for the DOD LPC-10 Vocoder in Quiet and E4B Acoustic Environments

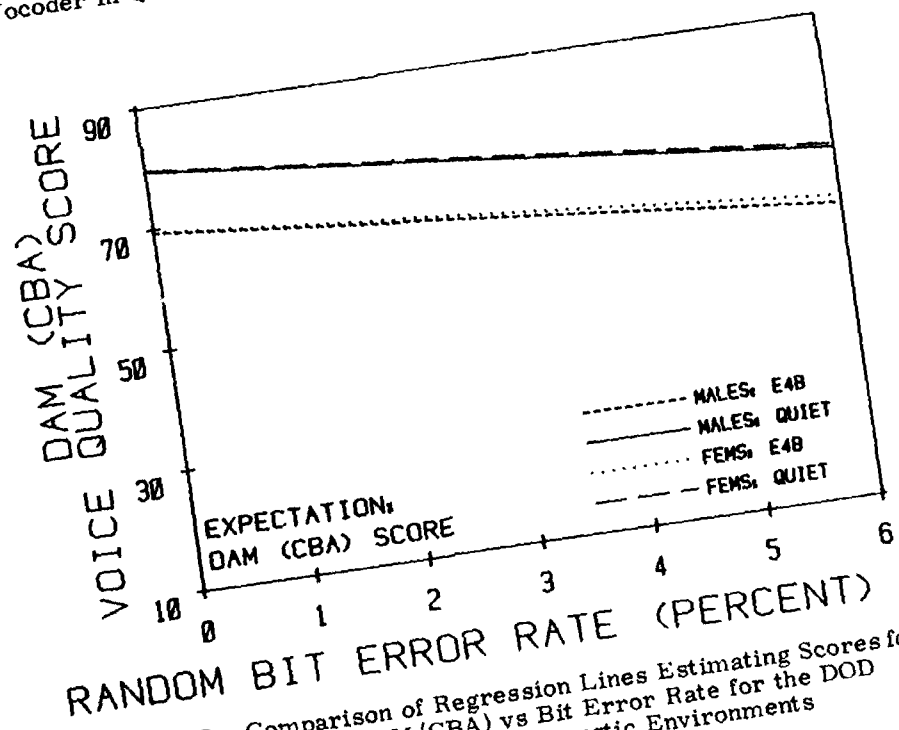


Figure 37. Comparison of Regression Lines Estimating Scores for Background Quality (DAM/CBA) vs Bit Error Rate for the DOD LPC-10 Vocoder in Quiet and E4B Acoustic Environments

12.1 Test Microphones

It has been observed that a noise cancelling microphone tends to produce slightly poorer intelligibility scores and voice quality scores in tests in a quiet environment than a high-quality dynamic microphone. However, this advantage quickly disappears when testing in all but the most benign noise environments because of the improvement in signal-to-noise ratio obtained with the noise cancelling microphone. The high quality microphone has been found to be the best choice for the ideal quiet conditions, but, for noise environments like those found on the E4B, it is likely that the Roanwell noise cancelling microphone used for these tests is at or very near to the optimum design.* This was not true of the microphone on the EC-135, as discussed in Section 12.7.

12.2 Test Speakers

Digital voice processors, especially narrowband processors, have been found to be very speaker-variable. Only very limited comparisons can be made of processors or test conditions unless speakers have been standardized. For tests of the effects of noise environments on vocoders and wideband voice processors, the DOD Digital Voice Processor Consortium has settled on a "standard" speaker set of three male and three female speakers. The Consortium speaker-set was used for these tests, expanded by adding three more male standard speakers to reduce experimental errors. Thus, the nine speakers used for these tests were a control set in the sense that the data on speech intelligibility and voice quality scores obtained with them is extensive enough to permit comparisons and statistical tests.

There is no assurance, however, that these speakers are a sample typical of command and control communicators for factors such as regional accents, age, voice pitch, etc. That uncertainty is mitigated by the fact that command and control communicators tend to have extensive experience in communicating under adverse conditions, able to adjust their voices to cope with noisy surroundings when speaking and also able to understand speech under adverse conditions when listening. Therefore, it is likely that "naive" speakers and listeners from a general civilian population would obtain results like those from these studies, while experienced military communicators would probably exceed these predictions. Further research is required to resolve this question.

---

* An experimental second-order noise cancelling microphone developed for overcoming severe acoustic noise environments was found to give poorer results than those presented here.

12.3 Random Bit Errors vs Block Errors

Considerable evidence exists to indicate that voice processor tests with random bit errors represent a "worst case," and that block errors, or clustered errors, that usually occur in transmitting digital data over wireline and radio channels have a smaller effect on intelligibility and voice quality than randomly distributed errors. From this point of view, the predicted effects of bit errors are conservative in the sense that data errors occurring in clusters, rather than in random distribution, will tend to degrade intelligibility and voice quality less than these forecasts.

12.4 Categories of Intelligibility and Voice Quality Performance

Categorizations of intelligibility and quality scores as good, poor, unacceptable, etc., represent a consensus reached by the DOD Digital Voice Processor Consortium from contacts with a number of workers in this field. These categorizations have not yet been validated by testing with military communicators. Scheduled research studies using conversational testing are expected to clarify the assignments of categories.

12.5 Results Not Applicable to LPC-10 Vocoders That Deviate From the DOD Standard Algorithm

The comparison illustrated in Figure 30 of Vocoder intelligibility with and without bit error correction under bit error conditions illustrates the fact that vocoders deviating from the DOD standard algorithm will probably obtain speech intelligibility and voice quality scores different from those presented here. The lack of the bit-error detection and correction features could significantly reduce performance in comparison with these values.

12.6 Comparison of Results From Live Recordings With Results From Simulations

Intelligibility scores obtained for the DOD LPC-10 vocoder with speech recordings prepared live aboard the aircraft during flight are compared with scores obtained from tests based on the simulation of the acoustic environment of the E4B Battlestaff Compartment in Figure 38. The cluster of scores and 95 percent confidence bounds for the set of data obtained with the simulation with speakers using the resident microphone are compared with the scores obtained with the live recordings when several microphones were used aboard the aircraft. None of the microphones gave any clear advantage over the Roanwell model used aboard the aircraft. Further details of the various microphones have been reported by Teacher.[5] The distribution of scores illustrates that live intelligibility tests prepared in the field tend to exhibit greater variability than do test recordings prepared under closely con-

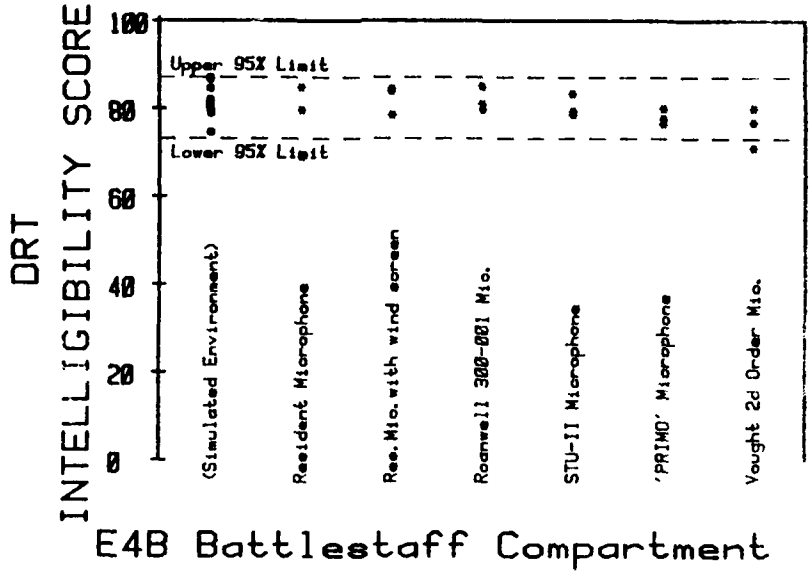


Figure 38. Comparison of Intelligibility Scores and 95% Confidence Bounds for Individual Speaker's Scores for the DOD LPC-10 Vocoder in the Simulated E4B Battlestaff Compartment Acoustic Environment With Scores From Live Recordings Aboard the Aircraft

trolled conditions. Even with these fluctuations, there is general agreement between results from the live recordings aboard the aircraft and the simulations of the acoustic environment.

12.7 Comparison of E4B and EC-135 Acoustic Environments

Vocoder intelligibility performance in the acoustic environment of the EC-135 was similar to that obtained for the E4B, probably because of the similarity in the background noise level and power spectrum. However, the voice quality (DAM/-CAE) scores obtained for the vocoder in the EC-135 acoustic environment were lower than scores obtained for the E4B. This is illustrated in Figure 39. Lower quality ratings were probably related to the differences in microphones; the speech-to-noise ratio provided by the pressure microphone used on the EC-135 was poorer than that ratio provided by the noise cancelling microphone used on the E4B.

13. CONCLUSIONS AND RECOMMENDATIONS

The E4B Battlestaff Compartment, NCA Compartment, and Briefing Room acoustic environments during flight were found to be not significantly different in

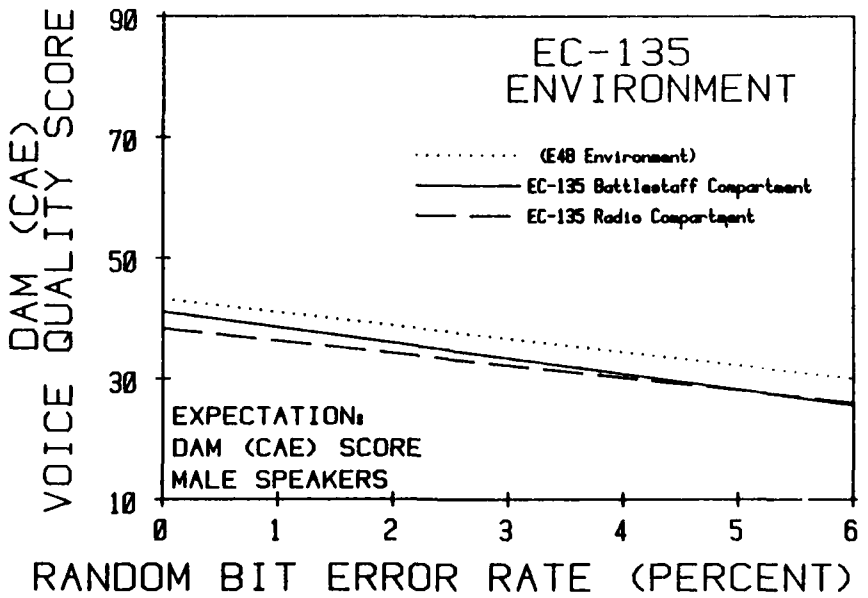


Figure 39. Comparison of Voice Quality (DAM/CAE) Scores vs Bit Error Rate for the DOD LPC-10 Vocoder in the EC-135 Acoustic Environment and in the E4B Composite Acoustic Environment

their effects on speaker's speech intelligibility and voice quality tested with the DOD standard LPC-10 narrowband vocoder.

Intelligibility and voice quality of the vocoder were significantly poorer with speakers in the acoustic environment of the E4B aircraft than with speakers in a quiet environment.

Tests with male and female speakers in the E4B acoustic environment indicated that average intelligibility of the vocoder operating with no data errors was 80.2 (fair), with 95 percent confidence bounds for individual speaker's intelligibility scores extending from 87.7 (good) to 72.7 (very poor). At 1 percent random bit-errors, average intelligibility was 75.7 (poor), with 95 percent confidence bounds for individual speaker's intelligibility extending from 83.2 (moderate) to 68.2 (unacceptable).

Average voice quality rating for the vocoder with speakers in the E4B acoustic environment was 43.3 (fair) in the absence of data errors, with 95 percent confidence bounds for individual speaker's quality ratings extending from 47.2 (fair) to 39.3 (poor). At a 2 percent error rate, the tests indicated an average voice quality rating of 38.7 (poor), with 95 percent confidence bounds for individual speaker's quality ratings extending from 42.6 (fair) to 34.7 (very poor).

At error rates of 1×10^{-3}, the intelligibility and quality of the vocoder would be essentially the intelligibility and quality quoted for the error-free condition.

Evaluation data indicated that, at error rates of 0.6 percent or less, 95 percent of the individual speakers from a population for which these were a representative sample would obtain intelligibility and quality above the unacceptable category.

Vocoders deviating from the DOD standard algorithm may not reach these performance levels, especially if they lack the bit-error detection and correction features.

A variation on LPC-10 vocoder design known as "piecewise-LPC" has design features that make it intrinsically less vulnerable to bit error effects and acoustic noise effects than a conventional LPC-10 vocoder.[13, 14] Our recommendations are that further research be done to complete the research in optimizing the PLPC vocoder design and incorporating bit error correction features, and that comparative evaluations be conducted to determine its effectiveness in alleviating the degradation of performance caused by noise and bit error effects.

---

13. Roberts, J.E., and Wiggins, R.H. (1976) Piecewise linear predictive coding (PLPC), Conf. Record ICASSP, IEEE Cat. No. 76CH1067-8 ASSP, 470-473.

14. Wiggins, R.H. (1976) Narrowband digital voice processing: Vol. II, CSP-30 PLPC software documentation, MTR-3324, MITRE Corp., Bedford, MA 01730, ESD-TR-76-282 Vol. II, AD B016323L.

References

1. Smith, C.P. (1977) Intelligibility Performance of Narrowband Linear Predictive Vocoders in the Presence of Bit Errors, AF Electronic Systems Division, Hanscom AFB, MA 01731, ESD-TR-77-328, AD A051323.
2. Smith, C.P. (1983) Relating the performance of speech processors to the bit error rate, Speech Tech. 2(No. 1).
3. Snedecor, G.W., and Cochran, W.G. (1967) Statistical Methods, Iowa State University Press, Ames, Iowa.
4. Tremain, T. (1982) The government standard linear predictive coding algorithm: LPC-10, Speech Tech. 1(No. 2).
5. Teacher, C.F. (1983) Secure Voice Evaluations: Characterization of Aircraft Noise and Audio Systems in the EC-135 and E4B Aircraft, Report KET-2709-2, Ketron, Inc., Wayne, Pa.
6. Voiers, W.D.; Cohen, M.E.; and Mickunas, J. (1965) Evaluation of Speech Processing Devices: 1. Intelligibility, Quality and Speaker Recognizability, AFCRL-65-826, Final Report, Contract AF19(628)-4195, AD 627320.
7. Voiers, W.D. (1977) Diagnostic evaluation of speech intelligibility, in Speech Intelligibility and Speaker Recognition, 2, M.E. Hawley, Ed., Benchmark Papers in Acoustics, Dowden, Hutchinson and Ross, Stroudsburg, Pa.
8. Voiers, W.D. (1983) Evaluating processed speech using the Diagnostic Rhyme Test, Speech Tech. 1(No. 3).
9. Voiers, W.D. (1977) Diagnostic acceptability measure for speech communications systems, IEEE Proc. ICASSP, 77CH1197-3 ASSP, 204-207.
10. Fairbanks, G. (1958) Test of phonemic differentiation: the Rhyme Test, J. Acoust. Soc. Am. 30:596-600.
11. House, A.S.; Williams, C.E.; Hecker, M.H.L.; and Kryter, K.D. (1965) Articulation testing methods: Consonantal differentiation with a closed response set, J. Acoust. Soc. Am. 37:158-166.

12. Preusse, J.W. (1959) The Consonant Recognition Test, U.S. Army Electronics Command, ECOM-3205, Fort Monmouth, N.J.

13. Roberts, J.E., and Wiggins, R.H. (1976) Piecewise linear predictive coding (PLPC), Conf. Record ICASSP, IEEE Cat. No. 76CH1067-8 ASSP, 470-473.

14. Wiggins, R.H. (1976) Narrowband digital voice processing: Vol. II, CSP-30 PLPC software documentation, MTR-3324, MITRE Corp., Bedford, MA 01730, ESD-TR-76-282 Vol. II, AD B016323L.

Appendix A

Intelligibility and Voice Quality Scores

Table A1. DRT Intelligibility Scores for the DOD LPC-10 Narrowband Vocoder in the E4B Battlestaff Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
(Test I.D.)	#9540	#9547	#9546	#9541	#9555	#9554	#9563	#9567	#9565	#9562
SPEAKERS										
Males:										
RH	84.6	86.7	83.1	83.7	79.6	79.9	73.0	71.6	63.7	67.3
JE	81.2	87.0	75.0	75.1	77.1	72.9	58.1	65.2	57.4	50.1
CH	81.6	80.7	80.5	76.0	74.0	72.0	68.7	74.7	51.8	61.7
CT	78.9	74.6	73.7	75.1	73.3	74.2	70.1	71.4	48.3	56.5
AS	80.6	79.7	77.6	76.0	71.0	77.5	69.4	70.4	52.7	55.1
BV	79.9	82.0	76.6	75.9	76.0	75.8	65.0	68.6	59.1	60.4
Females:										
VW	80.6	76.2	73.8	77.7	74.1	74.7	67.8	70.6	50.9	55.2
KS	84.0	83.7	78.4	82.8	73.4	75.7	74.9	76.7	63.3	64.6
MP	86.3	91.4	83.2	79.6	80.6	78.5	69.7	78.1	54.7	63.7
9-Speaker Avg	82.0	82.5	78.0	78.0	75.4	75.7	68.5	71.9	55.8	59.4

Table A2. DRT Intelligibility Scores for the DOD LPC-10 Narrowband Vocoder in the E4B NCA Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
(Test I.D.)	#9556	#9564	#9548	#9557	#9542	#9549	#9553	#9543	#9566	#9552
SPEAKERS										
Males:										
RH	84.9	88.7	81.6	81.4	77.6	77.3	70.1	70.6	63.4	66.0
JE	77.5	76.6	72.3	74.7	69.8	71.2	62.6	57.8	47.8	52.6
CH	87.6	81.6	82.9	79.0	82.3	78.0	72.1	66.8	62.2	61.6
CT	71.7	76.8	66.7	77.3	71.0	78.8	70.1	60.0	61.7	58.7
AS	76.2	80.3	78.5	72.8	73.4	76.8	69.5	68.2	58.3	60.7
BV	79.2	81.4	81.1	78.5	68.4	76.0	64.6	69.9	62.5	58.6
Females:										
VW	79.4	78.8	71.2	79.3	73.0	72.5	70.2	71.0	66.9	60.7
KS	83.3	88.0	82.0	80.6	77.3	80.2	78.1	68.5	53.5	62.0
MP	82.0	84.9	80.2	81.5	72.4	81.4	78.5	67.1	63.5	67.6
9-Speaker Avg	80.2	81.9	77.4	78.4	73.9	76.9	70.6	66.7	60.0	60.9

Table A3. DRT Intelligibility Scores for the DOD LPC-10 Narrowband Vocoder in the E4B Briefing Room Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
(Test I.D.)	#9560	#9569	#9568	#9559	#9558	#9561	#9550	#9545	#9544	#9551
SPEAKERS *Males:*										
RH	82.3	84.6	84.1	82.6	86.3	81.8	79.8	71.1	58.6	56.1
JE	82.9	78.0	79.4	80.6	75.4	79.0	69.3	63.9	55.3	58.7
CH	82.8	84.6	77.7	83.2	79.4	76.0	75.3	68.2	56.8	57.8
CT	77.2	75.8	79.2	75.3	72.8	76.0	69.0	65.8	62.4	52.6
AS	77.2	71.0	74.1	66.9	73.8	73.2	66.0	62.9	50.5	54.7
BV	81.5	77.7	78.0	77.1	76.3	72.8	69.1	65.9	59.5	52.0
6-Speaker Avg	80.7	78.6	78.3	77.6	77.3	76.5	71.4	66.3	57.2	55.3

Table A4. DAM/CAE (Overall Composite Voice Quality Score) for the DOD LPC-10 Vocoder in the E4B Battlestaff Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
DAM Test No.	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203
SPEAKERS										
Male:										
RH	45.3	48.5	46.4	43.7	44.0	41.2	40.7	41.9	32.9	32.5
JE	41.3	44.2	38.9	40.4	38.7	39.4	36.6	35.3	30.3	33.4
CH	39.7	41.9	36.0	40.5	38.8	37.3	37.9	39.0	31.3	31.9
CT	-	42.1	-	40.2	-	37.5	-	36.7	-	31.1
AS	-	45.4	-	42.6	-	41.5	-	41.4	-	36.2
BV	-	49.0	-	46.1	-	41.1	-	39.9	-	33.3
(Avg)	(42.1)	(45.2	(40.3)	(42.3)	(40.5)	(39.7)	(38.4)	(39.0)	(31.5)	(33.1)
Female:										
VW	40.5	-	40.1	-	39.6	-	37.4	-	30.0	-
KS	41.5	-	38.5	-	34.6	-	33.0	-	31.8	-
MP	47.8	-	43.6	-	41.7	-	37.6	-	33.3	-
(Avg)	(43.3)		(40.7)		(38.6)		(36.0)		(31.7)	

Table A5. DAM/CAE (Overall Composite Voice Quality Score) for the DOD LPC-10 Vocoder in the E4B NCA Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
DAM Test No.:	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203
SPEAKERS Males:										
RH	46.2	46.3	45.8	43.2	41.4	44.3	38.5	42.0	34.0	31.9
JE	43.1	44.6	36.3	41.9	41.4	42.4	36.8	38.4	31.8	31.7
CH	43.1	41.4	39.4	38.8	37.9	37.1	37.2	37.7	29.2	30.9
CT	-	46.1	-	44.2	-	41.5	-	36.5	-	31.5
AS	-	45.9	-	36.2	-	40.1	-	40.1	-	32.6
BV	-	44.8	-	45.1	-	40.8	-	42.2	-	35.4
(Avg)	(44.1)	(44.9)	(40.5)	(41.6)	(40.2)	(41.0)	(37.5)	(39.5)	(31.7)	(32.3)
Females:										
VW	44.9	-	43.2	-	42.9	-	41.1	-	33.1	-
KS	44.4	-	41.4	-	39.5	-	37.5	-	32.3	-
MP	45.8	-	43.8	-	41.1	-	39.3	-	30.6	-
(Avg)	(45.0)		(42.8)		(41.2)		(39.3)		(32.0)	

Table A6. DAM/CAE (Overall Composite Voice Quality Score) for the DOD LPC-10 Vocoder in the E4B Priefing Room Acoustic Environment (Fesident Microphone)

Bit Error Rate:	Zero	1/2% BER	1% BER	2% BER	5% BER
DAM Test No.:	D-203	D-203	D-203	D-203	D-203
SPEAKERS Males:					
RH	47.6	43.7	44.2	40.1	31.1
JE	43.0	39.2	40.5	38.1	30.8
CH	41.8	39.7	39.7	35.4	32.7
CT	42.2	43.8	39.9	37.6	31.9
AS	42.1	42.3	39.6	36.4	29.3
BV	44.0	41.2	42.8	39.9	29.0
(Avg)	(43.5)	(41.7)	(41.2)	(37.9)	(30.8)

Table A7. DAM/CSA (Signal Quality Scores) for the DOD LPC-10 Narrowband Vocoder in the E4B Battlestaff Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
DAM Test No.:	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203
SPEAKERS										
Males:										
RH	55.2	54.4	56.1	48.4	52.2	48.4	46.9	46.6	39.6	35.4
JE	53.3	55.1	46.9	49.5	47.8	48.9	43.7	44.9	38.3	39.8
CH	51.5	50.6	44.7	46.2	48.6	45.3	43.7	45.6	40.0	40.2
CT	-	51.4	-	49.2	-	42.5	-	42.5	-	36.2
AS	-	53.2	-	50.6	-	50.4	-	46.4	-	41.0
BV	-	57.9	-	55.8	-	52.4	-	48.8	-	41.7
(Avg)	(53.3)	(53.8)	(49.2)	(50.0)	(49.5)	(48.0)	(44.8)	(45.8)	(39.3)	(39.1)
Females:										
VW	51.7	-	53.4	-	50.2	-	48.0	-	38.8	-
KS	54.4	-	49.1	-	46.3	-	42.5	-	39.6	-
MP	56.2	-	52.9	-	50.5	-	47.4	-	42.4	-
(Avg)	(54.1)	-	(51.8)	-	(49.0)	-	(46.0)	-	(40.3)	-

Table A8. DAM/CBA (Background Quality Scores) for the DOD LPC-10 Narrowband Vocoder in the E4B Battlestaff Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
DAM Test No.:	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203
SPEAKERS										
Males:										
RH	74.9	72.9	74.8	70.7	71.9	67.2	76.1	67.3	61.9	59.2
JE	66.3	66.3	61.7	61.0	62.7	63.1	61.6	59.4	61.9	59.2
CH	64.4	65.9	64.2	67.1	63.8	63.9	62.6	64.6	61.3	60.2
CT	-	65.5	-	64.3	-	59.4	-	61.1	-	58.0
AS	-	72.4	-	73.3	-	69.6	-	68.4	-	67.1
BV	-	72.2	-	74.1	-	70.0	-	69.5	-	65.8
(Avg)	(68.5)	(69.2)	(66.9)	(68.4)	(66.1)	(65.5)	(66.8)	(65.1)	(61.7)	(61.6)
Females:										
VW	64.7	-	64.8	-	63.1	-	63.5	-	56.5	-
KS	61.6	-	60.3	-	63.5	-	60.0	-	59.1	-
MP	71.6	-	68.4	-	68.0	-	67.7	-	62.7	-
(Avg)	(66.0)	-	(64.5)	-	(64.9)	-	(63.7)	-	(59.4)	-

Table A9. DAM/CSA (Signal Quality Scores) for the DOD LPC-10 Narrowband Vocoder in the E4B NCA Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
DAM Test No.:	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203
SPEAKERS										
Males:										
RH	53.1	54.3	53.1	48.2	48.5	49.6	43.2	46.7	43.0	36.5
JE	52.7	51.9	49.0	49.7	47.5	50.8	46.7	46.7	37.5	37.7
CH	52.5	49.7	49.9	47.6	48.0	43.6	48.7	43.7	35.6	36.4
CT	-	53.2	-	50.0	-	48.8	-	42.6	-	39.0
AS	-	54.2	-	41.0	-	48.9	-	48.8	-	40.1
BV	-	53.3	-	52.7	-	48.2	-	49.7	-	41.5
(Avg)	(52.8)	(52.8)	(50.7)	(48.2)	(48.0)	(48.3)	(46.2)	(46.4)	(38.7)	(38.5)
Females:										
VW	54.7	-	50.7	-	50.4	-	48.5	-	42.7	-
KS	54.7	-	52.8	-	49.0	-	46.0	-	39.8	-
MP	53.1	-	51.9	-	50.3	-	46.0	-	39.3	-
(Avg)	(54.2)	-	(51.8)	-	(49.9)	-	(46.8)	-	(40.6)	-

Table A10. DAM/CBA (Background Quality Scores) for the DOD LPC-10 Narrowband Vocoder in the E4B NCA Compartment Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
DAM Test No.:	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203	D-202	D-203
SPEAKERS										
Males:										
RH	76.3	72.6	77.8	73.1	76.9	69.4	76.7	67.2	65.7	62.5
JE	68.1	70.3	62.3	68.9	63.4	66.8	64.7	63.3	57.5	55.9
CH	71.7	69.6	65.3	67.6	67.0	66.2	65.1	63.5	60.2	62.5
CT	-	71.9	-	73.4	-	66.4	-	66.5	-	61.0
AS	-	72.2	-	67.1	-	70.8	-	68.1	-	62.6
BV	-	74.5	-	73.2	-	70.8	-	71.6	-	69.8
(Avg)	(72.0)	(71.9)	(68.5)	(70.6)	(69.1)	(68.4)	(68.8)	(66.7)	(61.1)	(62.4)
Females:										
VW	73.3	-	72.7	-	70.7	-	69.1	-	63.5	-
KS	72.4	-	70.5	-	71.9	-	66.0	-	62.2	-
MP	75.8	-	72.2	-	71.2	-	69.8	-	64.5	-
(Avg)	(73.8)		(71.8)		(71.3)		(68.3)		(63.4)	

Table A11. DAM/CSA (Signal Quality Scores) for the DOD LPC-10 Narrowband Vocoder in the E4B Briefing Room Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero	1/2% BER	1% BER	2% BER	5% BER
DAM Test No.:	D-203	D-203	D-203	D-203	D-203
SPEAKERS					
Males:					
RH	54.7	47.0	49.6	46.6	34.9
JE	51.1	48.7	48.1	46.0	36.0
CH	50.1	48.9	48.7	44.0	37.7
CT	50.0	51.0	44.8	41.4	37.8
AS	49.4	50.0	46.7	40.6	36.4
BV	53.7	50.7	50.1	47.7	36.9
(Avg)	(51.5)	(49.4)	(48.0)	(44.4)	(36.6)

Table A12. DAM/CBA (Background Quality Scores) for the DOD LPC-10 Narrowband Vocoder in the E4B Briefing Room Acoustic Environment (Resident Microphone)

Bit Error Rate:	Zero	1/2% BER	1% BER	2% BER	5% BER
DAM Test No.:	D-203	D-203	D-203	D-203	D-203
SPEAKERS					
Males:					
RH	74.5	74.7	74.5	65.0	62.0
JE	64.0	64.4	64.5	60.9	60.3
CH	71.9	68.2	67.8	67.7	60.5
CT	70.7	70.7	68.8	69.2	64.4
AS	71.8	73.0	71.9	69.9	65.6
BV	74.0	75.8	72.7	72.3	74.6
(Avg)	(71.2)	(71.1)	(70.0)	(67.5)	(64.6)

Table A13. DRT Intelligibility Scores for the DOD LPC-10 Narrowband Vocoder in a Quiet Acoustic Environment (Dynamic Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
(Test I.D.)	#1406	#1408	#1443	#1447	#1445	#1448	#1442	#1449	#1444	#1446
SPEAKERS										
Males:										
RH	90.5	90.4	90.0	87.2	89.2	90.4	82.7	85.5	71.6	81.9
JE	88.4	87.5	80.6	86.2	85.0	81.2	82.8	80.3	78.5	75.1
CH	92.3	93.7	92.1	87.9	91.5	87.9	87.6	84.9	78.5	83.3
LL	90.2	88.5	91.7	91.9	91.0	86.6	83.2	85.2	77.2	82.8
BV	88.4	90.2	88.0	83.7	86.6	86.7	85.4	81.9	81.0	80.1
PK	88.3	88.2	88.0	87.5	90.1	85.8	82.8	82.4	77.0	78.6
(Avg)	(89.7)	(89.8)	(88.4)	(87.4)	(88.9)	(86.4)	(84.1)	(83.4)	(77.3)	(80.3)
Females:										
VW	89.1	89.7	84.1	82.6	86.7	83.1	80.2	81.2	81.2	78.6
KS	79.8	82.3	86.1	80.7	83.3	84.6	81.6	84.8	72.3	76.3
MP	87.2	87.9	85.0	86.3	85.3	77.0	81.8	78.6	79.0	79.8
JS	87.5	88.0	82.8	77.5	86.6	78.5	76.7	82.0	66.3	67.2
LS	80.7	81.1	80.2	80.3	82.2	81.0	77.3	75.9	77.1	69.7
LV	85.0	84.6	80.5	83.5	82.7	79.7	78.6	79.0	70.6	71.0
(Avg)	(84.9)	(85.6)	(80.5)	(83.5)	(82.7)	(79.7)	(78.6)	(79.0)	(74.4)	(73.8)

Table A14. DAM/CAE (Overall Composite Voice Quality Score) for the DOD LPC-10 Narrowband Vocoder in a Quiet Acoustic Environment (Dynamic Microphone) From Three Independent DAM Tests

Bit Error Rate:	Zero			1/2% BER			1/% BER			2% BER			5% BER		
DAM Test Nos.:	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106
SPEAKERS															
Males:															
RH	49.2	48.0	49.7	47.1	48.1	44.7	45.6	44.5	42.1	43.9	43.3	38.8	36.6	35.7	31.2
JE	50.8	49.7	48.6	47.8	47.5	43.4	43.1	43.5	42.8	42.1	39.5	39.5	32.3	34.9	31.9
CH	46.3	47.4	48.1	47.1	42.9	42.5	45.5	44.5	39.0	43.8	41.3	36.9	36.2	34.7	32.4
(Avg)	48.8	48.4	48.8	47.3	46.2	43.5	44.7	44.2	41.3	43.3	41.4	38.4	35.0	35.1	31.8
Females:															
VW	43.4	45.7	43.7	45.4	44.6	39.7	43.7	43.5	38.7	41.5	38.7	34.0	34.2	31.8	31.8
KS	45.8	45.0	44.7	46.8	45.3	40.4	45.2	42.3	38.9	40.9	39.1	37.2	33.7	30.6	31.4
MP	46.2	46.7	44.6	44.8	45.7	41.0	42.2	44.9	39.7	39.9	39.3	37.8	32.6	33.8	33.4
(Avg)	45.1	45.8	44.3	45.7	45.2	40.4	43.7	43.6	39.1	40.8	39.0	36.3	33.5	32.1	32.2

Table A15. DAM/CSA (Signal Quality Scores) for the DOD LPC-10 Narrowband Vocoder in a Quiet Acoustic Environment (Dynamic Microphone) From Three Independent DAM Tests

Bit Error Rate:	Zero			1/2% BER			1% BER			2% BER			5% BER		
DAM Test Nos.:	D-101	D-101	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106
SPEAKERS															
Males:															
RH	63.8	61.7	64.7	62.4	62.4	61.7	60.6	60.9	61.5	59.4	61.1	55.1	51.9	52.7	49.3
JE	63.7	65.4	63.6	62.2	61.8	61.7	60.4	58.4	62.1	53.6	51.1	58.1	45.6	49.0	50.1
CH	60.9	62.3	64.1	59.0	56.5	62.2	57.0	59.3	59.2	56.5	55.0	57.1	50.8	46.7	49.7
(Avg)	61.8	63.1	64.1	61.2	60.2	61.9	59.3	59.5	60.9	56.5	55.7	56.8	49.4	49.5	49.7
Females:															
VW	57.3	60.6	59.4	61.4	59.1	56.8	59.0	59.1	57.6	57.8	55.1	51.5	48.9	43.8	46.9
KS	59.3	59.0	59.2	60.5	59.4	57.6	59.7	56.0	57.5	54.9	53.2	54.2	51.0	44.6	49.0
MP	58.8	60.3	60.7	59.3	60.2	56.9	54.1	61.0	59.5	53.7	53.3	53.7	44.6	46.6	50.1
(Avg)	58.5	60.0	59.8	60.4	59.6	57.1	57.6	58.7	58.2	55.5	53.9	53.1	48.2	45.0	48.7

Table A16. DAM/CBA (Background Quality Scores) for the DOD LPC-10 Narrowband Vocoder in a Quiet Acoustic Environment (Dynamic Microphone) From Three Independent DAM Tests

Bit Error Rate:	Zero			1/2% BER			1% BER			2% BER			5% BER		
DAM Test Nos.:	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106
SPEAKERS															
Males:															
RH	78.4	80.0	81.6	78.3	82.4	74.3	78.9	78.5	73.7	76.6	74.8	74.1	70.1	67.8	65.1
JE	79.6	81.0	82.9	82.1	79.9	77.6	82.0	76.9	75.7	80.7	79.6	70.8	78.2	72.8	70.0
CH	78.3	81.4	80.7	78.5	75.9	74.8	78.2	77.6	70.5	76.6	76.8	68.4	71.7	69.7	64.0
(Avg)	78.8	80.8	81.7	79.7	79.5	75.6	79.8	77.7	73.4	78.0	77.1	71.2	73.4	70.2	66.3
Females:															
VW	79.3	81.5	79.8	79.1	80.5	78.0	78.5	78.2	74.0	77.2	74.9	72.0	74.1	72.9	68.1
KS	80.8	80.6	84.9	81.7	80.4	75.6	81.4	76.0	75.9	77.8	77.1	73.4	71.9	69.1	70.5
MP	76.2	80.3	81.4	78.7	79.7	75.2	78.1	81.1	71.8	75.4	74.4	74.2	73.3	7.13	63.9
(Avg)	78.8	80.8	82.0	79.8	80.2	76.3	79.3	78.4	73.9	76.8	75.5	73.2	73.1	71.1	67.5

Table A17. DRT Intelligibility Scores for the CVSD-16 Processor in a Quiet Acoustic Environment (Dynamic Microphone) From Three Independent Intelligibility Tests

Bit Error Rate:	Zerb			1/2% BER			1% BER			2% BER			5% BER		
(Test I.D.)	#1346	#1404	#1017	#1345	#1383	#1580	#1344	#1382	#1582	#1343	#1380	#1576	#1342	#1381	#1578
SPEAKERS															
Males:															
RH	92.7	95.7	93.4	94.4	95.6	91.9	92.2	95.1	92.4	92.7	94.0	93.4	90.0	91.9	88.9
JE	88.4	92.6	90.0	92.6	92.2	92.2	91.1	87.5	91.4	88.0	92.1	90.4	80.7	87.4	82.8
CH	90.5	94.8	95.4	91.8	96.1	91.5	95.1	95.1	93.6	90.8	95.7	95.4	88.5	90.9	90.2
(Avg)	90.5	94.4	92.9	92.9	94.6	91.9	92.8	92.6	92.5	90.5	93.9	93.1	86.4	90.1	87.3
Females:															
VS	92.3	90.2	89.6	87.0	87.6	91.4	87.6	89.3	89.2	82.6	84.4	88.0	81.1	85.8	81.4
KS	88.0	91.0	89.3	87.6	88.7	86.6	89.6	88.2	90.0	85.5	89.6	85.9	86.3	85.8	84.1
MP	90.0	90.9	92.3	85.3	87.5	91.3	85.0	87.4	89.6	85.0	87.5	85.3	84.5	86.7	80.3
(Avg)	90.1	90.7	90.4	86.6	87.9	89.8	87.4	88.3	89.6	84.4	87.2	86.4	84.0	86.1	81.9

Table A18. DAM/CAE (Overall Composite Voice Quality Score) for the CVSD-16 Processor in a Quiet Acoustic Environment (Dynamic Microphone) From Three Independent DAM Tests

Bit Error Rate:	Zero			1/2% BER			1% BER			2% BER			5% BER		
DAM Test Nos.:	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106
SPEAKERS															
Males:															
RH	52.0	57.4	51.2	45.3	50.2	45.7	46.5	49.8	46.1	47.1	48.7	43.8	41.6	44.0	38.7
JE	59.6	64.7	53.3	50.8	52.5	49.8	49.6	50.2	45.5	50.8	51.0	43.4	47.4	48.4	37.8
CH	55.6	54.6	51.4	48.5	51.6	49.8	47.2	50.6	48.2	49.3	48.9	45.0	43.6	45.6	34.7
(Avg)	55.7	58.9	52.0	48.2	51.4	48.4	47.8	50.2	46.6	49.1	49.5	44.1	44.2	46.0	37.1
Females:															
VW	47.0	51.0	51.3	45.4	49.2	50.3	44.8	47.8	46.5	45.4	46.5	44.0	41.4	44.1	36.1
KS	52.9	57.3	49.3	47.7	49.8	45.8	47.7	50.1	45.8	47.9	48.4	40.6	43.8	45.6	35.4
MP	53.5	57.7	51.3	45.7	51.1	44.9	45.2	49.8	45.2	46.3	50.1	43.3	42.2	47.2	38.3
(Avg)	51.1	55.3	50.6	46.3	50.0	47.0	45.9	49.2	45.8	46.5	48.3	42.6	42.5	45.6	36.6

Table A19. DAM/CSA (Signal Quality Scores) for the CVSD-16 Processor in a Quiet Acoustic Environment (Dynamic Microphone) From Three Independent DAM Tests

Bit Error Rate:	Zero			1/2% BER			1% BER			2% BER			5% BER		
DAM Test Nos.:	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-1	D-101	D-105	D-106
SPEAKERS															
Males:															
RH	67.6	71.3	73.4	65.7	69.9	67.5	68.7	70.7	68.1	67.0	69.1	64.2	64.9	66.2	57.7
JE	71.5	74.9	74.3	72.0	73.5	71.2	71.8	72.3	66.0	72.0	72.5	62.6	69.7	70.3	60.4
CH	68.4	68.0	73.6	69.8	70.5	74.1	69.6	69.7	68.7	70.5	68.4	64.9	67.2	67.2	52.7
(Avg)	69.2	71.4	73.8	69.2	71.3	70.9	70.0	70.9	67.6	69.8	70.0	63.9	67.3	67.9	56.9
Females:															
VW	63.0	67.0	70.8	67.3	68.8	71.5	66.2	67.4	69.2	66.2	68.7	65.6	65.1	67.0	57.9
KS	65.6	69.9	66.7	65.4	68.6	66.4	67.3	68.6	64.9	64.6	67.5	60.6	67.0	66.3	57.6
MP	67.8	68.6	72.7	65.4	70.7	66.1	66.8	68.6	67.9	64.1	71.1	63.5	62.1	68.1	58.1
(Avg)	65.4	69.4	70.0	66.0	69.4	68.0	66.8	68.2	67.3	65.0	69.1	63.2	64.7	67.1	57.9

Table A20. DAM/CBA (Background Quality Scores) for the CVSD-16 Processor in a Quiet Acoustic Environment (Dynamic Microphone) From Three Independent DAM Tests

Bit Error Rate:	Zero			1/2% BER			1% BER			2% BER			5% BER		
DAM Test Nos.:	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106	D-101	D-105	D-106
SPEAKERS															
Males:															
RH	82.5	86.2	72.8	68.9	76.2	73.0	70.9	74.6	70.5	72.6	75.1	69.9	67.1	68.7	66.4
JE	84.9	86.7	72.4	73.2	77.2	74.3	72.0	75.3	71.6	74.4	76.8	68.5	70.6	73.0	65.1
CH	82.9	85.4	70.6	71.4	77.1	73.2	70.3	77.4	72.1	71.8	76.2	71.5	67.8	71.0	64.1
(Avg)	83.4	86.1	71.9	71.2	76.8	73.5	71.1	75.8	71.5	72.9	76.0	70.0	68.5	70.9	65.2
Females:															
VW	84.1	87.0	73.5	70.5	75.8	75.5	70.1	73.7	73.7	70.3	73.1	71.3	66.2	69.2	63.1
KS	85.5	85.5	76.0	74.3	77.8	79.7	75.0	77.4	75.1	74.1	76.5	71.7	71.1	72.7	66.4
MP	82.7	85.6	70.7	70.3	77.6	73.2	70.8	75.0	71.8	71.3	75.8	70.8	67.5	73.1	68.1
(Avg)	84.1	86.0	73.5	71.7	77.0	76.2	72.0	75.4	73.6	71.9	75.2	71.3	68.3	71.7	65.9

Table A21. DRT Intelligibility Scores for the CVSD-32 Processor in a Quiet Acoustic Environment (Dynamic Microphone)

Bit Error Rate:	Zero	1/2% BER	1% BER	2% BER	5% BER
(Test I.D.)	#1405	#1320	#1316	#1319	#1317
SPEAKERS					
Males:					
RH	95.8	96.4	94.1	95.2	92.6
JE	93.2	92.4	95.1	94.0	90.5
CH	96.5	95.2	96.1	94.0	94.7
(Avg)	(95.2)	(94.7)	(95.1)	(94.4)	(92.6)
Females:					
VW	92.6	91.7	93.6	86.3	82.8
KS	94.7	91.8	89.6	90.5	89.7
MP	90.1	89.2	92.1	89.1	89.5
(Avg)	(92.5)	(90.9)	(91.8)	(88.6)	(87.3)

Table A22. DAM/CAE (Overall Composite Voice Quality Score) for the CVSD-32 Processor in a Quiet Acoustic Environment (Dynamic Microphone)

Bit Error Rate:	Zero		1/2% BER		1% BER		2% BER		5% BER	
DAM Test No.:	D-101	D-105	D-101	D-105	D-101	D-105	D-101	D-105	D-101	D-105
SPEAKERS										
Males:										
RH	62.9	67.2	53.4	58.2	54.6	56.8	51.3	54.4	45.6	49.1
JE	65.9	71.3	54.0	58.9	54.6	57.8	54.5	57.9	50.4	50.3
CH	63.8	62.9	55.0	56.3	51.7	56.4	52.4	51.6	49.0	46.7
(Avg)	(64.2)	(67.1)	(54.1)	(57.8)	(53.6)	(57.0)	(52.7)	(54.6)	(48.3)	(48.7)
Females:										
VW	55.4	58.6	49.1	54.2	38.4	54.0	49.2	49.6	45.7	44.9
KS	62.6	66.9	52.5	55.9	52.3	56.4	52.6	53.2	46.6	46.3
MP	57.6	63.7	50.7	57.3	49.0	55.3	50.1	54.7	44.4	48.5
(Avg)	(58.6)	(63.1)	(50.8)	(55.8)	(46.6)	(55.2)	(50.6)	(52.5)	(45.6)	(46.6)

Appendix B

List of Master Tapes for Evaluating
Speech System Performance in the
E4B and EC-135 Acoustic Environments

Table B1. List of Master Tapes: E4B Battlestaff Compartment Acoustic Environment (Roanwell 240100001 N/C Microphone)

K10-1.1-A*	RH 323A	K10-1.1-B*	CT 301B	K10-1.1-C*	VW 335A
	JE 314B		AS 306B		KS 315B
	CH 309B		BV 303A		MP 331A
K10-1.2-A	RH 323B	K10-1.2-B	CT 301A	K10-1.2-C	VW 335B
	JE 315A		AS 305A		KS 316B
K10-1.3-A	RH 324A	K10-1.3-B	CT 302A	K10-1.2-C	VW 336A
	JE 315B		AS 306A		KS 316A
	CH 310A		BV 303B		MP 332B
K10-1.4-A	RH 324B	K10-1.4-B	CT 302B	K10-1.4-C	VW 336B
	JE 314A		AS 305B		KS 315A
	CH 310B		BV 304A		MP 331B

*This tape includes the Three-Speaker DAM following the last DRT.

Table B2. List of Master Tapes: E4B NCA Compartment Acoustic Environment (Roanwell Model 240100001 Noise-Cancelling Microphone)

K10-2.1-A*	RH 325A	K10-2.1B*	CT 308B	K10-2.1-C*	VW 301A
	JE 312A		AS 312A		KS 317B
	CH 308A		BV 309A		MP 333A
K10-2.2-A	RH 326A	K10-2.2-B	CT 307B	K10-2.2-C	VW 302A
	JE 312B		AS 311A		KS 318B
	CH 308B		BV 309B		MP 334A
K10-2.3-A	RH 326B	K10-2.3-B	CT 308A	K10-2.3-C	VW 302B
	JE 313B		AS 312B		KS 318A
	CH 307B		BV 310A		MP 334B
K10-2.4-A	RN 325B	K10-2.4-B	CT 307A	K10-2.4-C	VW 301B
	JE 313A		AS 311B		KS 317A
	CH 307A		BV 310B		MP 333B

*This tape includes the Three-Speaker DAM following the last DRT.

Table B3. List of Master Tapes: E4B Briefing Room Acoustic Environment (Roanwell Model 240100001 Noise-Cancelling Microphone)

K10-3.1-A*	RH 327B	K10-3.1-B*	CT 313A
	JE 317A		AS 329B
	CH 305A		BV 315A
K10-3.2-A	RH 327B	K10-3.2B	CT 313B
	JE 318A		AS 330B
	CH 305B		BV 315B
K10-3.3-A	RH 328A	K10-3.3-B	CT 314A
	JE 318B		AS 330A
	CH 306B		BV 316A
K10-3.4-A	RH 327A	K10-3.4-B	CT 314B
	JE 317B		AS 329A
	CH 306A		BV 316B

*This tape includes the Three-Speaker DAM following the last DRT.

Table B4. List of Master Tapes: EC-135 Battlestaff Compartment Acoustic Environment (Roanwell Model 60150 Microphone)

K11-1.1-A*	RH 330A	K11-1.1-B*	CT 317A	K11-1.1-C*	VW 303A
	JE 309B		AS 321B		KS 311B
	CH 304A		BV 323A		MP 320A
K11-1.2-A	RH 329B	K11-1.2-B	CT 317B	K11-1.2-C	VW 303B
	JE 309A		AS 322B		KS 312B
	CH 303B		BV 324A		MP 319A
K11-1.3-A	RH 329A	K11-1.3-B	CT 318A	K11-1.3-C	VW 304B
	JE 310A		AS 322A		KS 311A
	CH 304B		BV 324B		MP 320B
K11-1.4-A	RH 330B	K11-1.4-B	CT 318B	K11-1.4-C	VW 304A
	JE 310B		AS 321A		KS 312A
	CH 303A		BV 323B		MP 319B

*This tape includes the Three-Speaker DAM following the last DRT.

Table B5. List of Master Tapes: EC-135 Radio Compartment Acoustic Environment (Roanwell Model 60150 Microphone)

K11-2.1-A*	RH 331A	K11-2.1-B*	CT 319A	K11-2.1-C*	VW 305A
	JE 307B		AS 328B		KS 314A
	CH 301B		BV 326A		MP 321B
K11-2.2-A	RH 331B	K11-2.2-B	CT 320B	K11-2.2-C	VW 306B
	JE 308B		AS 327B		KS 313B
	CH 302B		BV 325B		MP 322B
K11-2.3-A	RH 332A	K11-2.3-B	CT 320A	K11-2.3-C	VW 306A
	JE 308A		AS 327A		KS 313A
	CH 301A		BV 326B		MP 321A
K11-2.4-A	RH 332B	K11-2.4-B	CT 319B	K11-2.4-C	VW 305B
	JE 307A		AS 328A		KS 314B
	CH 302A		BV 325A		MP 322A

*This tape includes the Three-Speaker DAM following the last DRT.

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.

Printed by
United States Air Force
Hanscom AFB, Mass. 01731