MONTANA
STATE UNIVERSITY

Speech-signal processing and applications to single-sideband
by Norman A Shyne

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in Electrical Engineering
Montana State University
© Copyright by Norman A Shyne (1962)

Abstract:
The results of signal processing as applied to SSB systems is described in this paper.

In the preliminary phases of the research, certain procedures, designs, and definitions are established: 1. An articulation test procedure 2. The definition of a measure of signal-to-noise ratio useful for peak-limited systems 3. General equipment and experiment designs as related to SSB signal processing A review of processing as applied to audio signals is presented and the equivalence of intelligibility scores for fixed- and variable-volume listening conditions is cited. Linear processes of power-spectrum equalization and differentiation are shown to provide minor intelligibility increases for λ defined at audio. For λ defined at narrow-band and in conjunction with finite degrees of clipping at audio, DSB, and SSB these linear operations generally give rise to intelligibility degradation. Intelligibility considerations for an audio AVC unit are presented; the results indicate that, under certain conditions, the AVC unit provided nearly the same intelligibility increases as 12 db of audio clipping.

The three most fundamental (and accurate) intelligibility estimates were made for: audio, DSB, and SSB clipping with λ defined at narrow-band. Of these, SSB clipping provided the most significant increase (for a given signal peak value) and at the same time more fully preserved signal quality. It is shown, for example, that for a given noise level, 60% intelligibility can be maintained by using an SSB signal which has been clipped 36 db and has a peak value 12 db below that of an unprocessed signal. The interdependent and iterative combinations of the three clipping operations yielded advancements in intelligibility estimates; however, in no case did they exceed those derived from an equivalent amount of SSB clipping.

The final consideration of this investigation was a reduced-X signal condition and its effect upon the intelligibility estimates for audio and SSB clipping. For both audio and SSB, the reduced-X signal yielded lower intelligibility scores. It appears that a parametric relationship of the form C max= λ+δ,&dealta;≈ db exists for maximum clip- ping; greater amounts of clipping decrease intelligibility scores for a given amount of noise.

SPEECH-SIGNAL PROCESSING AND APPLICATIONS TO SINGLE-SIDEBAND

by

N. A. Shyne

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

Approved:

Byron J. Barnett
Head, Major Department

Donald K. Weaver
Chairman, Examining Committee

Leon H. Johnson
Dean, Graduate Division

MONTANA STATE COLLEGE
Bozeman, Montana

March, 1962

Copyright © by N. A. Shyne

1962

ACKNOWLEDGEMENTS

I wish to express my appreciation to Professor Donald K. Weaver, Jr. for his thoughtful guidance and encouragement in this research. Thanks are also extended to Mr. John F. Bowker for his part in the realization of the experimental design, compilation of test materials, and suggested improvements of the manuscript.

The research reported in this paper has been sponsored by the Electronics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF19(604)7365.

TABLE OF CONTENTS

		Page
	ABSTRACT	x
1.	INTRODUCTION	1
	1.1. History and Background	1
	1.2. Signal Processing Related to Communication Systems	2
	1.3. A Statement of the Research Problems	4
2.	DEFINITIONS, DERIVATIONS, AND THEORETICAL CONSIDERATIONS	7
	2.1. Properties of Speech Signals	7
	2.2. Speech Transmission Systems	8
	2.2.1. SSB: Filter Method	9
	2.2.2. Linearity of Frequency Translation	10
	2.3. Signal-to-Noise Ratio	11
	2.4. Signal Processing	13
	2.4.1. Applicability	13
	2.4.2. Processing Devices	14
	2.4.2.1. Clipper	14
	2.4.2.2. Linear Filters	16
	2.4.2.3. Automatic Volume Control	20
3.	EQUIPMENT AND EXPERIMENT DESIGN	22
	3.1. System	22
	3.2. The Articulation Test	24
	3.2.1. Equipment	24
	3.2.2. Test Materials	25
	3.2.3. Test Team	26
	3.3. Tape Preparation and Sample Articulation Test	30
	3.4. Signal Tableaux	35
4.	DATA PRESENTATION AND INTERPRETATION	40
	4.1. Data Reduction Sample	40
	4.2. Test Results	42
	4.2.1. Preliminary Study	43
	4.2.2. Audio Processes for SSB	48

Page

4.2.3. Narrow-band Processes. 53
4.2.4. Iterative, Interdependent and Other Combinatorial Processes 58

4.2.4.1. Iterative Processes. 58
4.2.4.2. Interdependent Processes 59
4.2.4.3. Linear Pre-Processing. 62
4.2.4.4. Reduced Signal λ 63

5. DISCUSSION; INTERPRETATIVE LIMITATIONS 66

6. RESULTS AND CONJECTURE 69

7. APPENDICES 72

Appendix A - Literature Consulted. 73

Appendix B - Signal Types. 75

Appendix C - Model SSB Systems 76

1. Filter Method. 76
2. QF Method. 76

Appendix D - Processing Devices. 82

1. Clipper. 82
2. Mathematical Considerations of Clipping. 84
3. Integrator, Differentiator, AVC, Equalizer 90

Appendix E - Additional System and Measurement Considerations. 95

1. Earphone Analysis. 95
2. True rms vs. Average-reading rms 97
3. λ Change Due to Clipping 98
4. Bandwidth and λ. 100
5. Channel λ. 101

Appendix F - Tabular Data. 103

Appendix G - Tests of Significance 107

LIST OF ILLUSTRATIONS

Figure		Page
2-1.	An Oscillogram of the Word "add"	7
2-2.	Filter Method of SSB Generation	9
2-3.	Generalized Communication System	13
2-4.	Idealized Clipper Characteristic	15
2-5.	Peak Clipper in Block Notation	16
2-6.	SSB, DSB, Audio Filter Characteristics	17
2-7.	SSB Clipper Followed by a 3 kc Mechanical Filter	18
2-8.	Electrical Integrator	19
2-9.	Amplitude Characteristics of the Integrator and Differentiator	20
3-1.	Filter Method of SSB Generation and Processing	23
3-2.	Spectral Studies of the SSB System (No Processing)	23
3-3.	Relative Peak Values of Words of List 74	27
3-4.	Peak List Words for Lists 1-100	27
3-5.	Audiograms of "typical" and the "worst" Members Selected for the Test Team	28
3-6.	Control Scores for the First Five Test Sessions for an Articulation Test Team	28
3-7.	A Typical Subject Response Sheet	31
3-8.	Typical Test Score Sheet	31
3-9.	Representative Recording Set-up for n db of SSB Clipping, No Other Processing	32
3-10.	Peaking (ΔPV) Due to Filtering After SSB Clipping	34
3-11.	Oscillograms of the Word "Poem-Fate" for Various System Signals	36
3-12	Sonagrams of the Word "Poem-Fate" for Various System Signals	38
4-1.	Frequency Polygon for $\mathbb{P}C^{s}_{24}F^{s}_{m}$, $\lambda^{s} = 9$	41
4-2.	Frequency Polygon for $\mathbb{P}C^{s}_{12}F^{s}_{m}$, $\lambda^{s} = 3$	41
4-3.	Intelligibility Data for Initial Control, λ^{a}	44
4-4.	Modified Control - λ^{a} Presentation	44
4-5.	E and $T_f E$	44

Figure		Page
4-6.	Sample Study of Audio Clipping, λ^a	44
4-7.	Intelligibility Test of Audio AVC for Maximum Input Voltage of 0.5 v p-p, λ^a	47
4-8.	Audio AVC for Constant Decay-Time of 0.5 sec, λ^a	47
4-9.	Combinations of V and C_n^a, λ^a	47
4-10.	Intelligibility Estimates for Audio Clipping at 3 kc Bandwidth, λ^a	47
4-11.	Intelligibility Estimates for $\mathbb{P}C_n^aF_m^s$ (λ^s)	50
4-12.	I vs $\mathbb{P}C_n^dF_m^s$ (λ^s)	52
4-13.	I vs $\mathbb{P}C_n^sF_m^s$ (λ^s)	54
4-14.	Intelligibility Estimates for Iterative SSB Clipping: A Comparison	57
4-15.	A Comparison of Single and Iterative Audio Clipping for SSB, $\lambda^s = 15$	57
4-16.	Interdependencies Between Audio and DSB Clipping, $\lambda^s = 15$	57
4-17.	Interdependencies Between Audio and SSB Clipping, $\lambda^s = 9$	57
4-18.	Interdependencies Between Audio and SSB Clipping, $\lambda^s = 15$	61
4-19.	The Effect of Audio Pre-Clipping on $C_{36}^sF_m^s$	61
4-20.	The Effect of T_r Pre-Processing on $C_n^aF_m^s$	61
4-21.	The Effect of T_r and E Pre-Processing on C_n^s	61
4-22.	The Effect of T_r on $C_6^sF_m^sC_{36}^sF_m^s$	64
4-23.	The Effect Upon I of Reducing Signal λ Prior to Audio Clipping, λ^s	64
4-24.	The Effect Upon I of Reducing Signal λ Prior to SSB Clipping, λ^s	64
B-1.	Band-limited Signals	75
C-1.	Laboratory SSB System Components	77
C-2.	Quadrature-Function Method of SSB Generation	78
C-3.	QF SSB Generator	80
C-4.	Spectral Study of QF Generator	80
D-1.	Theoretical Clipper Considerations	83

Figure		Page
D-2.	High Frequency Clipper	84
D-3.	Spectral Component Groups for an nth-Order Approximation to a Clipping Function	88
D-4.	Circuit Diagram of the Differentiator	91
D-5.	Circuit Diagram of the Integrator	91
D-6.	Circuit Diagram of the AVC	92
D-7.	Input-Output Characteristics of the AVC	92
D-8.	Circuit Diagram of the Speech Equalizer	94
D-9.	Speech Equalizer Response Characteristic	94
E-1.	Equipment Configuration for Threshold Hearing Tests	96
E-2.	Response of Earphone, Position #2	96

LIST OF TABLES

Table		Page
F-1.	Data for $\mathbb{P}C^{s}_{24}F^{s}_{m}$, $\lambda^{s} = 9$	104
F-2.	Data for $\mathbb{P}C^{s}_{12}F^{s}_{m}$, $\lambda^{s} = 3$	104
F-3.	Data for Initial Control, λ^{a}	104
F-4.	Data for Modified Control, λ^{a}	104
F-5.	Data for E and $T_{f}E$, λ^{a}	104
F-6.	Data for Audio Clipping Sample, λ^{a}	104
F-7.	Data for Audio AVC: Maximum Input Voltage of 0.5 v p-p	104
F-8.	Data for Audio AVC: Constant Decay-Time of 0.5 sec	104
F-9.	Data for Combinations of V and C^{a}_{n}, λ^{a}	104
F-10.	Data for $C^{a}_{n}F$ (3 kc Bandwidth), λ^{a}	105
F-11.	Data for $C^{a}_{n}F^{s}_{m}$, λ^{s}	105
F-12.	Data for $\mathbb{P}C^{d}_{n}F^{s}_{m}$, λ^{s}	105
F-13.	Data for $\mathbb{P}C^{s}_{n}F^{s}_{m}$, λ^{s}	105
F-14.	Data for $\mathbb{P}C^{s}_{n}F^{s}C^{s}_{n}F^{s}_{m}$ and $\mathbb{P}C^{s}_{n}F^{s}_{t}C^{s}_{n}F^{s}_{m}$	105
F-15.	Data for $C^{a}_{n}FC^{s}_{n}F^{s}_{m}$, $\lambda^{s} = 15$	105
F-16.	Data for Interdependencies: Audio and DSB, $\lambda^{s} = 15$	105
F-17.	Data for Interdependencies: Audio and SSB, $\lambda^{s} = 9$	106
F-18.	Data for Interdependencies: Audio and SSB, $\lambda^{s} = 15$	106
F-19.	Data for $\mathbb{P}C^{a}_{n}F^{s}C^{s}_{36}F^{s}_{m}$	106
F-20.	Data for $T_{r}\ C^{a}_{n}F^{s}_{m}$	106
F-21.	Data for $T_{r}\ C^{s}_{n}F^{s}_{m}$ and $E\ C^{s}_{n}F^{s}_{m}$	106
F-22.	Data for $T_{r}\ C^{a}_{6}F^{s}_{m}C^{s}_{36}F^{s}_{m}$	106
F-23.	Data for $\lambda^{27}\ C^{a}_{n}F^{s}_{m}$, λ^{s}	106
F-24.	Data for $\lambda^{27}\ C^{s}_{n}F^{s}_{m}$, λ^{s}	106

ABSTRACT

The results of signal processing as applied to SSB systems is described in this paper.

In the preliminary phases of the research, certain procedures, designs, and definitions are established:

1. An articulation test procedure
2. The definition of a measure of signal-to-noise ratio useful for peak-limited systems
3. General equipment and experiment designs as related to SSB signal processing

A review of processing as applied to audio signals is presented and the equivalence of intelligibility scores for fixed- and variable-volume listening conditions is cited. Linear processes of power-spectrum equalization and differentiation are shown to provide minor intelligibility increases for λ defined at audio. For λ defined at narrow-band and in conjunction with finite degrees of clipping at audio, DSB, and SSB these linear operations generally give rise to intelligibility degradation. Intelligibility considerations for an audio AVC unit are presented; the results indicate that, under certain conditions, the AVC unit provided nearly the same intelligibility increases as 12 db of audio clipping.

The three most fundamental (and accurate) intelligibility estimates were made for: audio, DSB, and SSB clipping with λ defined at narrow-band. Of these, SSB clipping provided the most significant increase (for a given signal peak value) and at the same time more fully preserved signal quality. It is shown, for example, that for a given noise level, 60% intelligibility can be maintained by using an SSB signal which has been clipped 36 db and has a peak value 12 db below that of an unprocessed signal. The interdependent and iterative combinations of the three clipping operations yielded advancements in intelligibility estimates; however, in no case did they exceed those derived from an equivalent amount of SSB clipping.

The final consideration of this investigation was a reduced-λ signal condition and its effect upon the intelligibility estimates for audio and SSB clipping. For both audio and SSB, the reduced-λ signal yielded lower intelligibility scores. It appears that a parametric relationship of the form $C_{max} = \lambda + \delta$, $\delta \approx 8$ db exists for maximum clipping; greater amounts of clipping decrease intelligibility scores for a given amount of noise.

SPEECH-SIGNAL PROCESSING AND APPLICATIONS TO SINGLE-SIDEBAND

1. INTRODUCTION

1.1. History and Background

The process of communication is fundamental to the existence of our society as we know it today. As used in the general sense, the term communication is broadly inclusive; humans communicate with one another in myriad ways. In the case of human-to-human communications, the spoken word is probably the most important vehicle of information, although the communicative process may be carried on by appealing to any of the senses.

For many years researchers have placed a firm concentration on auditory speech signals; and, as a result, have gained an immeasurable amount of information concerning the speech production and speech detection mechanisms. Initially, such information was of primary concern to psychologists and linguists. However, with the advent of rapidly expanding telephone networks in the 1920's, engineers became aware of the problems of transmitting aural signals and turned to previous knowledge. For example, in 1922 Stewart proposed "An Electrical Analog of the Vocal Organs"; Crandall and MacKenzie (1922) analyzed the energy distribution of speech. Later, Steinberg (1929) and Knudsen (1929) considered the effects of echoes, phase distortion, and reverberation on the intelligibility of speech. Martin (1930) initiated bandwidth studies; Dudley's "The Vocoder" (1939) embodied fundamental concepts which are in wide use in present bandwidth-reduction schemes.

World War II provided a great impetus to research investigations concerned with speech intelligibility. Many significant results were obtained during this period. Of particular interest to the war effort was, of course, the improvement of intelligibility of speech in the presence of noise for military applications. Some contributors during this era were Egan et al (1943), Egan and Wiener (1946), Smith (1946), French and Steinberg (1947), Pollack (1948), and Licklider and Pollack

(1948). These investigations confirmed that a speech signal and the characteristics of the human ear are such as to tolerate remarkable amounts of alteration or processing of the signal without giving undue loss of intelligibility.

Since 1950, the acquisition of knowledge concerning speech production, processing, transmission, and reception has been directed in five general areas:

1) Signal processing related to communication systems
2) Bandwidth conservation
3) Oral machine control and machine translation of languages
4) Equipments
5) Linguistics, speech, psychology

The first four areas have been of interest to engineers and scientists in the general field of communication. Research in the second, third, and fourth areas has been typified by advancements in time and frequency domain vocoders, digital-computer simulation of speech (David, Mathews, MacDonald, 1958), auditory recognition (David, 1958), and complex systems for military and civilian communications networks. The first area, of particular interest to the author, has acquired additions such as speech peak clipping (Martin, 1951; Kahn, 1957; Wathen-Dunn and Lipke, 1958); the evaluation of processing devices and more sophisticated system evaluation techniques.

1.2. Signal Processing Related to Communication Systems

Communication systems range in complexity from a simple conversation between two humans standing face-to-face to the most highly complex electronic devices covering a span of thousands of miles. Even with this latitude of complexity, a speech communications system is generally construed to be composed of a transmitter, a receiver, and an inter-connecting channel. System requirements then determine the complexity of each of the system components.

Due to its high ratio of peak-to-average values, a speech signal is not generally well-suited for transmission over a given communication

system. The question arises then: "Can a certain form of processing make a speech signal more suitable for transmission?" Fortunately, researchers have recognized this problem and have found that speech signals possess certain elements which are not essential to the psychoacoustic process of speech recognition. However, no one has been able to show precisely the totality of what elements are essential or why. On a piecewise basis it has been shown that certain bands of frequencies in the spectrum of a speech signal are not required for recognition; it has also been shown that a severe restriction of peak excursions of a signal will not seriously impair its intelligibility. So without real definition, researchers have undertaken the task of mapping auditory speech signals to a domain where they are manageable to the point of (hopefully) preserving only the information-bearing characteristics. Circuit and system components have been designed to preserve these characteristics, with constraints applied by the ultimate receptor of the signal.

In the communicative process, the receptor of a speech signal is a human being (barring a few possibilities such as voice-controlled machines). If a signal is to be useful, it must be intelligible; signal intelligibility is a relative measure of the effectiveness of the transmission of the signal with a given set of system parameters. Unfortunately, there exists no instantaneously-reading meter which displays intelligibility as a function of signal-to-noise ratio, speaker clarity and enunciation, system bandwidth, signal envelope preservation, or the listener's ability to hear. The only available meter, if it may be called such, for the determination of single-word articulation scores is a test using a team of listeners. Certain limitations and perturbations of speech signals are observed by a team of normal, well-trained listeners. From this point, one possibility is for the listeners to write down what they hear; these written results compared to a standard key will define articulation (or intelligibility) score.

Recent conditions of overcrowding the spectrum space have prompted numerous investigations of bandwidth compression techniques. Perhaps due to the concern regarding this spectrum crowding, certain aspects of signal processing as applied to analog speech signals in communication systems have been slighted or perhaps even overlooked. The application of more sophisticated combinations of speech processing techniques does not seem to have been generally accepted or widely recognized as a valid possibility of improving system performance. On the contrary, most present military specifications for voice communications equipment specify rather rigid tolerances on linear response and a uniform speech pass-band, though some include provisions for a nominal amount of peak clipping. This situation is somewhat surprising in view of the possible improvements in system performance which might be obtained at nominal costs through the application of more sophisticated processing techniques.

1.3. A statement of the Research Problems

Signal processing is a mapping operation. Presumably in an actual system this mapping must be performed by devices which retain at least part of the information-bearing characteristics of the signals. For the case of single-sideband (SSB) communication systems, there has been no concerted effort to determine: what degree of processing may be performed, or what effect processing in one part of the system has upon signal characteristics in another part.

As a result of definition and physical restrictions, signal processing may be categorized into the following types:

I. Linear
II. Instantaneous non-linear
III. Slow operations on averages
IV. Fast operations

A process of P(V) is linear if for every pair of variables V_1 and V_2 and for each pair of constants k_1 and k_2 the following relation holds:

$$P(k_1V_1 + k_2V_2) = k_1 P(V_1) + k_2 P(V_2) \qquad (1\text{-}1)$$

Examples of linear processing are: spectrum tilts, frequency translation, and equalization of the power spectrum. Type II processing requires that there exist an output signal which is uniquely defined for each input signal and is independent of the past history of the input (no memory); peak clipping is an example of such processing. Type III effects changes in signal characteristics as dictated in some unique way by the mean or average of the signal; an example is automatic gain control. Type IV processing is contingent upon fast operations which are dependent upon the harmonic content of the signal. Note that each of these processes may be considered an analog signal operation: the input and output signals of these processing devices are describable by equations of similar form on a one-to-one basis. The vocoder, by this definition, is not an analog operation.

For purposes of this investigation, the better of two processes will be taken as that which yields the highest intelligibility score for given signal bandwidth and noise corruption. Further criteria such as "quality" and "naturalness" of the processed signal as presented to the receptor will not be of primary concern.

Obviously, it was impossible to investigate all combinations of processes applicable to all systems. This study was delimited to an investigation of various linear and non-linear operations on analog speech signals in a representative, peak-limited, SSB system; the intrinsic effect of these operations upon the psycho-acoustic process of speech recognition was of primary concern. Only combinations of Types I, II, and III processing were considered.

In order to study signal processing as applied to SSB systems, it is necessary to define certain signal and system models to which

fundamental constraints may be applied. To effect this end, basic definitions, derivations, assumptions, and theoretical considerations are outlined in Chapter 2. Most of this information is not new, but it forms a logical basis for the experimental design presented in Chapter 3. The results of the study are given in Chapter 4, and a discussion based upon these results is presented in Chapters 5 and 6. In general, circuit diagrams and descriptions, lengthy derivations, and subsidiary data are placed in the appendixes following the literature consulted.

2. DEFINITIONS, DERIVATIONS, AND THEORETICAL CONSIDERATIONS

Signals and systems may generally be classified into major groups, with relatively precise group definitions. In order to clarify divers discussions of signal processing as applied to SSB systems that follow, certain classifications, definitions and assumptions are necessary.

2.1. Properties of Speech Signals

As initially derived from the human voice-production mechanism, speech is a continuous, single-valued function of time having a dc component.* A speech signal is limited in frequency range and considered to be a broadband** signal. The peak-factor (defined as the ratio of the peak to rms values) of a speech signal is about 15 to 20 db. In general, the peak excursions of the signals are unsymmetrical (even

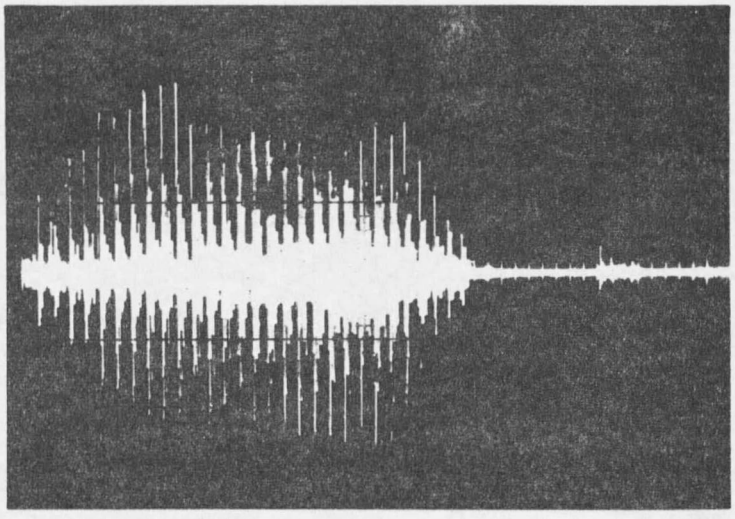


Figure 2-1. An Oscillogram of the word "add" band-limited from 300 to 3000 cps.

---

* The "breath" bias due to exhaling while speaking normally.
**Defined in Appendix B.

if the dc component is neglected) and a sample of finite duration contains a finite power. These properties become evident after consideration of the "time-domain picture" of the word "add" as shown in Figure 2-1. It is convenient for the communications engineer to consider an analog speech signal as an electrical waveform, representative of the sound-pressure waveform, with a frequency characteristic restricted to the baseband (typically 300 to 3000 cps). Obviously, the analog signal no longer contains the dc component.

2.2. Speech Transmission Systems

The process of modulation is an inherent part of signal transmission. A modulation may be effected in a variety of ways which yield a variety of signal types. Some familiar types of modulation are:

1. Amplitude
2. Double-sideband, suppressed carrier (DSB or D)
3. Single-sideband (SSB or S)
4. Phase
5. Frequency
6. Pulse Code

With due regard to the attendant problems of other methods of modulation, the delimitation to consider only an SSB system was made for the purposes of this investigation; however, some of the results obtained may be applicable to other methods.

The ideal process of generating an SSB signal is a linear frequency-translation to the domain of a narrow-band signal. In practice, translation may be effected by one of the following methods:

1. Filter
2. Phase
3. Quadrature Function (QF)

The first two methods have been recognized and used extensively; Weaver (1958) has pointed out advantages of the QF method. The decision to use the filter method of SSB generation for this study was made after

considerable time and effort were expended in the design of a QF generator. Although more degrees of freedom of signal processing are possible using the QF method, tone and unwanted-sideband suppression were inadequate for purposes of a predictable laboratory system. Circuit diagrams and a comparison of QF and filter system responses are given in Appendix C.

2.2.1. SSB: Filter Method

The basic SSB generator, shown in the block notation of Figure 2-2,

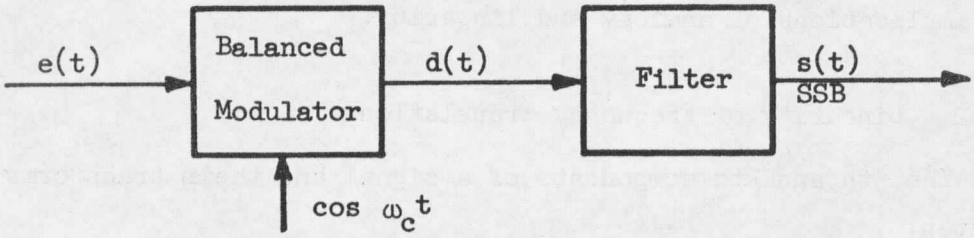


Figure 2-2. Filter Method of SSB Generation

may be analyzed in the following manner. Assume a series representation for the input signal. The input is then given by

$$e(t) = \Sigma\, a_i \cos(\omega_i t + \phi_i). \tag{2-1}$$

The balanced modulator forms the product:

$$d(t) = e(t) \cos \omega_c t = \cos \omega_c t \left[\Sigma\, a_i \cos(\omega_i t + \phi_i)\right]$$

$$d(t) = \tfrac{1}{2} \Sigma\, a_i \left[\cos(\omega_c t + \omega_i t + \phi_i) + \cos(\omega_c t - \omega_i t + \phi_i)\right] \tag{2-2}$$

The two cosine functions of (2-2) are the sidebands of a DSB signal. Assuming the filter selects the upper sideband,

$$s(t) = \tfrac{1}{2} \Sigma\, a_i \cos\,(\omega_c t + \omega_i t + \phi_i). \tag{2-3}$$

Note that (2-3) has the same form as (2-1); hence s(t) is the analog of e(t) under a change of variable $\omega_i' = \omega_c + \omega_i$. Note that each ω_i was linearly translated in frequency to occupy the spectral line $(\omega_i + \omega_c)$. In the actual case, an SSB filter (for example a Collins mechanical filter, characterized in Figure 2-6) will not have a uniform phase characteristic and the ϕ_i will be modified. This does not change the implications of analogy and linearity.

2.2.2. Linearity of Frequency Translation

The jth and kth components of a signal and their transforms can be written:

$$e_j(t) = a_j \cos\,(\omega_j t + \phi_j)$$

$$e_k(t) = a_k \cos\,(\omega_k t + \phi_k)$$

$$s_j(t) = T\,[e_j(t)] = C\, a_j \cos\,(\omega_o t + \omega_j t + \phi_j)$$

$$s_k(t) = T\,[e_k(t)] = C\, a_k \cos\,(\omega_o t + \omega_k t + \phi_k)$$

Form: $$T[e_j(t) + e_k(t)] = \tag{2-4}$$

$$C[a_j \cos\,(\omega_o t + \omega_j t + \phi_j) + a_k \cos\,(\omega_o t + \omega_k t + \phi_i)] =$$

$$T[e_j(t)] + T[e_k(t)] \tag{2-5}$$

But (2-4) and (2-5) are true for all j and k; hence, by (1-1) frequency translation is a linear process (transformation).

Frequency translation is a completely reversible process. The power spectrum and amplitude function* of a signal are invariant with frequency translation and uniform phase shift (Weaver, 1958). The envelope and amplitude function are identical for a narrow-band signal, whereas the envelope of a broadband signal has little meaning.

2.3. Signal-To-Noise Ratio

The term "signal-to-noise ratio", per se, will not be used in this report. In order to avoid possible confusions, a variable λ is defined:

$$\lambda = 20 \log_{10} \frac{E_p}{E_n} , \qquad (2\text{-}6)$$

where E_p is the signal peak value and E_n is the rms value of the corrupting white noise; E_p and E_n are of equal bandwidth. Since the prime consideration of this study was peak-limited systems, the standard definition of signal-to-noise ratio was unacceptable. A peak-power constraint of a system implies a peak limitation of the narrow-band modulating function. Since there is no fundamental relationship between the peak and RMS values of a short-time speech sample, utilization of the classical S/N would have been ineffectual.

An interesting fact is that signal peak-value (PV) is variant with a frequency translation from baseband to narrowband; that is, ΔPV exists. An intuitive implication of ΔPV may be shown by an example. Signal frequency components which are harmonically related at baseband are of nearly the same frequency at narrowband. Consequently, it is impossible for

*Defined as the root of the sum of squares of the quadrature functions.

the spectral components to add in identically the same way; hence, the peak value must change with such a translation. In general, if PV changes, λ changes and $\Delta\lambda$ exists. Since ΔPV is a function of signal phase relationships, frequency translation, and signal processes (even in the linear case) it provides a useful control in the preparation of test materials and the evaluation of certain signal processes.

When linear processing and constant bandwidth are maintained, the changes in PV, λ, and PF* (expressed in db) are identical:

$$\Delta\lambda = \Delta PV = \Delta PF$$

This may be seen from the following:

$$\Delta\lambda = \lambda_1 - \lambda_2$$

$$= 20 \log \frac{E_{p_1}}{E_n} - 20 \log \frac{E_{p_2} k_2}{E_n k_2} \text{ , (where } k_2 \text{ is a gain constant)}$$

$$= 20 \log \frac{E_{p_1} E_n k_2 E_s}{E_{p_2} E_n k_2 E_s} = PV_1 - PV_2 = \Delta PV$$

$$= 20 \log \frac{E_{p_1}}{E_s} - 20 \log \frac{k_2 E_{p_2}}{k_2 E_s} = PF_1 - PF_2 = \Delta PF$$

$\therefore \quad \Delta\lambda = \Delta PF = \Delta PV$ (in db) Q. E. D.

---

*Peak factor (PF) is defined as $20 \log_{10} \frac{E_p}{E_s}$, where E_p and E_s are peak and rms signal values, respectively.

2.4. Signal Processing

2.4.1. Applicability

As applied to a generalized communication system shown in Figure 2-3, signal processing may be effected in one of the following ways:

1. Signal operations in the original spectrum space (baseband).
2. Translated signal operations

In the system of Figure 2-3, the baseband frequencies occur at both the

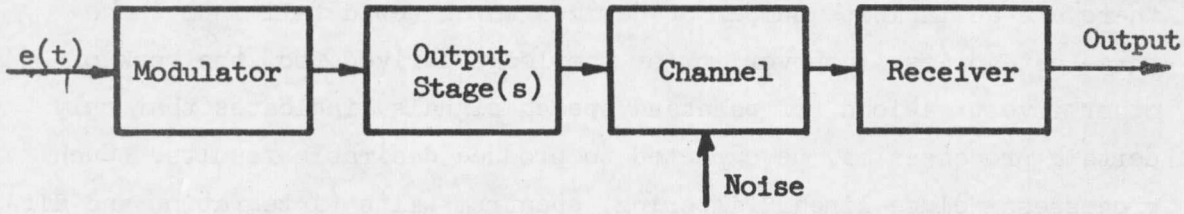


Figure 2-3. Generalized Communications System

input and the output. In many systems the modulator and output stage are combined and may contain numerous possibilities for the application of processing. Processing may in general be achieved at the input and output of the channel. For an SSB system, a reference to Figure 2-2 and equation (2-3) indicates that signal processing may be applied to:

1. Original audio signal
2. DSB prior to filtering
3. SSB
4. Receiver output

When a speech signal is considered a random function of time, the a_i of (2-1) through (2-3) are random variables. Variations of the a_i (or ϕ_i) may be effected by the utilization of a linear device whether the signal is in original or translated spectrum space. A prediction may be made as to the new a_i and ϕ_i from the amplitude and phase

characteristics of the linear device. For non-linear processing, however, there is no convenient method of mathematical analysis suitable for speech, and experimental investigation is warranted.

2.4.2. Processing Devices

With the possible points of application of signal processing in SSB communication systems enumerated, the question arises: what kind of processing at each of these points might produce a desirable result as concerns the system output signal? The general classifications of processing devices to be used were: linear, instantaneous non-linear, and slow operations on averages (types I, II and III processing). Theoretically there are an infinite number of devices which would fall into these three categories. However, prior knowledge derived from the work of other investigations (of baseband speech signals) indicates that only certain processes may be expected to produce desirable results. Such processes include linear filtering, spectrum tilts (integration and differentiation), clipping, and automatic volume control.

Only fundamental explanations of the basic processing devices will be made here. More detailed descriptions of analysis, design, and operation are presented in Appendix D.

2.4.2.1. Clipper

Licklider (1946) has pointed out that one of the fundamental information-bearing components of a speech signal is the temporal axis-crossing pattern. A peak-clipping device obviously does not change this pattern since, in the vicinity of the zero axis, the clipper is essentially a linear device. The peak clipper is an example of Type 2 processing, and an idealized transfer characteristic is given in Figure 2-4.*

*An analytical approximation to a clipper characteristic is given in Appendix D. Predictive spectrum-spreading considerations are also presented.

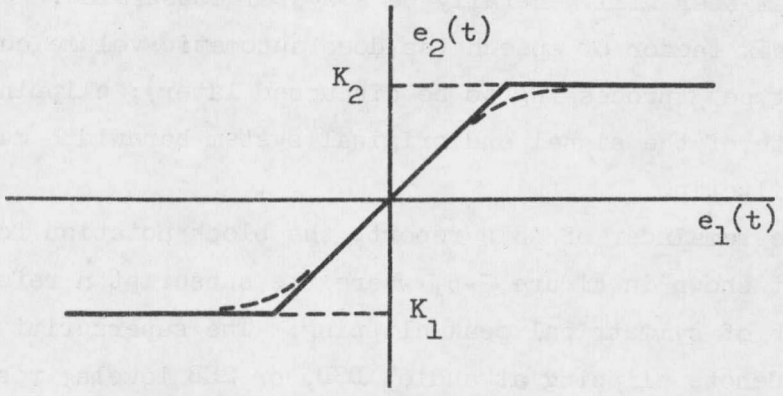


Figure 2-4. Idealized Clipper Characteristic

A clipper may be defined in terms of its terminal conditions:

$$e_2(t) = e_1(t) \text{ when } K_1 \leq e_1(t) \leq K_2$$

$$e_2(t) = K_1 \text{ when } e_1(t) \leq K_1$$

$$e_2(t) = K_2 \text{ when } e_1(t) \geq K_2 ,$$

where K_1 and K_2 are the predetermined saturation levels of the clipper. A clipper produces a distortion such that the amplitude of the signal never exceeds a given level. There is a limit, however, since extreme clipping and post-amplification can increase background noise level to the point where the speech signal is no longer useful.

Clipping, as an example of non-linear processing, has been discussed extensively in the literature. For example, curves in Wathen-Dunn and Lipke (1958) indicate that, for 12 db of audio peak-clipping, the carrier power of an AM communication transmitter could be reduced 7.6 db when clipping increased the sideband power by 8.5 db (for a given value of intelligibility). Hence, although intelligibility is a

prime measure of any communication system, the peak-power capability of the transmitter will generally be a system constraint. Clipping reduces the peak factor of speech (as does automatic volume control, an example of type 3 processing to be discussed later); clipping increases the bandwidth of the signal and original system bandwidth must be restored by filtering.

For the remainder of this report, the block notation for a clipper will be that shown in Figure 2-5, where the subscript n refers to the degree in db of symmetrical peak clipping. The superscript will be a, d, or s to denote clipping at audio, DSB, or SSB levels, respectively.

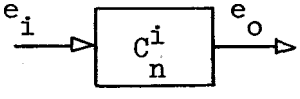
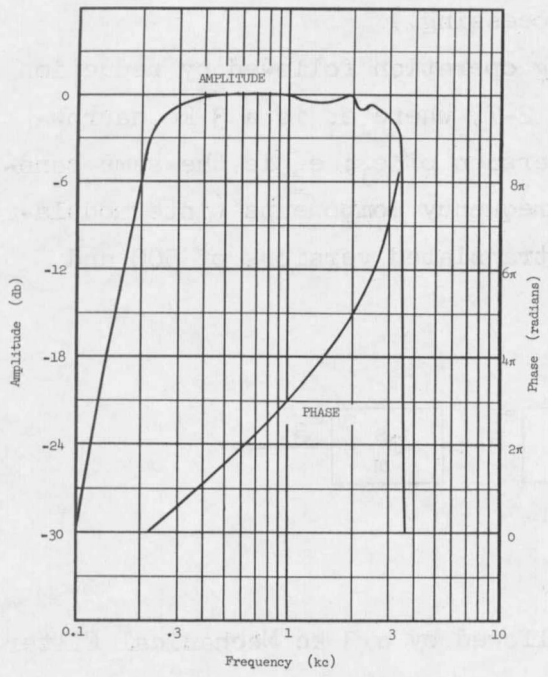


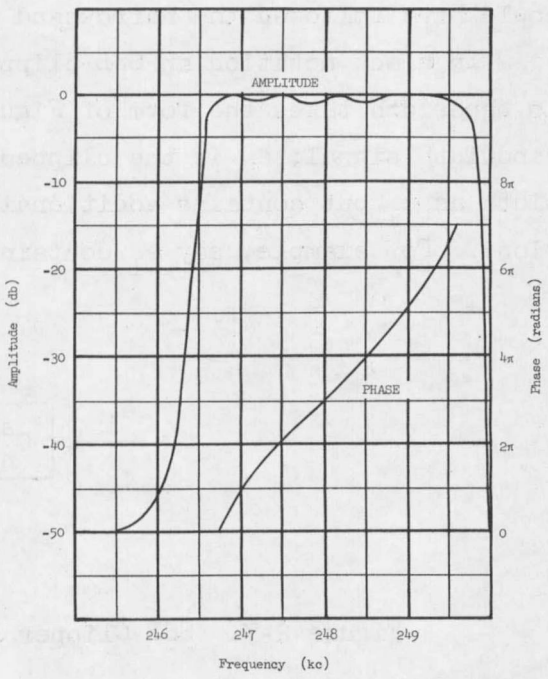
Figure 2-5. Peak Clipper in Block Notation

2.4.2.2. Linear Filters

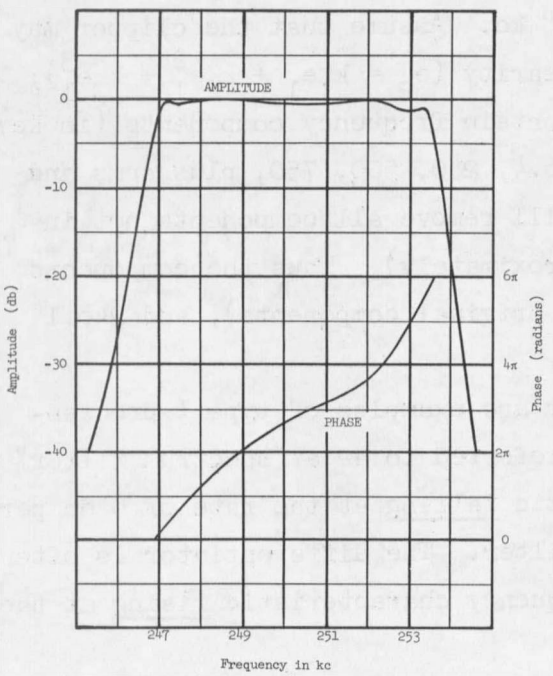
The symbol "F" will be used to denote a band-pass filtering operation at audio frequency. The actual filters used were 600-ohm, M-derived sections designed to give a 300 to 3000 cps bandwidth, closely approximating the bandwidth of Collins F 250-Z5 SSB filters. SSB mechanical filters will be designated F_m^s in order to distinguish them from tuned processing filters at narrow-band (F_t^s). A double-bandwidth mechanical filter, F_m^d, (Collins 250A67) was used to provide a double-bandwidth signal with adequate skirt rejection. Amplitude and phase characteristics of these filters are given in Figure 2-6. These particular filter designs were chosen as typical of filters encountered in actual SSB systems. (It is important to realize that a translation to the domain of a <u>different narrowband</u> signal would not change the implications of this study. This is true since the envelope and amplitude functions of a narrowband signal are invariant with frequency translation (Weaver,



a). Audio



b). SSB



c). DSB

Figure 2-6. Amplitude and Phase Characteristics of Filters:
a). Audio
b). SSB
c). DSB

1958). Consequently, 455 kc mechanical filters could have been used, assuming identical phase characteristics, or another translation to HF could have followed the narrowband processing.)

In block notation an SSB clipping operation followed by reduction to bandwidth takes the form of Figure 2-7, where e_1 is a 3 kc narrowbandwidth signal; e_2 is the clipped version of e_1; e_3 is the same bandwidth as e_1 but contains additional frequency components (intermodulation). For example, say e_1 contains translated versions of 500 and

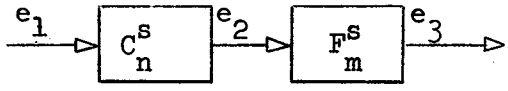


Figure 2-7. SSB Clipper Followed by a 3 kc Mechanical Filter

1200 cps: if the translation frequency is 250 kc, the lower sideband (e_1) has components 249.5 kc and 248.8 kc. Assume that the clipper may be approximated by a 3rd-order nonlinearity ($e_2 = k_1e_1 + k_2e_1^2 + k_3e_1^3$); the output of the clipper will then contain frequency components (in kc): 249.5, 248.8, 499.0, 497.6, 748.5, 746.4, 250, 500, 750, plus sums and differences. The mechanical filter will remove all components not in the desired range 247.0 to 249.7 (approximately). Thus the components of e_3 will be: 248.8, 249.5 (the two original components), and 248.1 (497.6 - 249.5).

The integrator and differentiator are examples of type I processing devices. An integrator is often referred to as a "spectrum tilter" with a frequency response characteristic <u>falling</u> at the rate of 6 db per octave; in essence it is a low-pass filter. The differentiator is often called a "spectrum tilter" with a frequency characteristic <u>rising</u> at the

rate of 6 db per octave. Figure 2-8 gives a schematic diagram for a simplified electrical integrator.

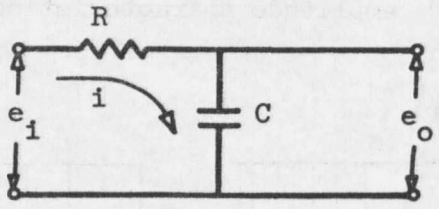


Figure 2-8. Electrical Integrator

If $V_c = 0$ at $t = 0$, then

$$I(s) = \frac{E_i(s)}{R + \frac{1}{Cs}},$$

where $E_i(s)$ and $I(s)$ are the transforms to the complex frequency domain of $e_i(t)$ and $i(t)$, respectively.

$$E_o(s) = \frac{I(s)}{Cs} = \frac{E_i(s)}{1 + sRC}.$$

If $sRC \gg 1$, $RC \gg \frac{1}{s} = \frac{1}{j\omega} = \frac{T}{j2\pi}$ $(T = \frac{1}{f})$,

then $$E_o(s) = \frac{E_i(s)}{sRC} = \frac{E_i(j\omega)}{j\omega RC}. \qquad (2\text{-}7)$$

For a value of RC large with respect to the longest time increment of the signal, the output voltage is proportional to the integral of the input signal. (2-7) also indicates that, for a fixed E_i (hence e_i) and an octave increase in frequency, E_o decreases by a factor of 2 (or 6 db). It is on this basis that an integrator is often referred to as a spectrum "tilter".

An electrical differentiator may be considered the dual of the circuit shown in Figure 2-8 with a characterizing analysis analogous to the foregoing. Circuit diagrams of the actual integrator and differentiator are given in Appendix D; amplitude characteristics of these devices are presented in Figure 2-9.

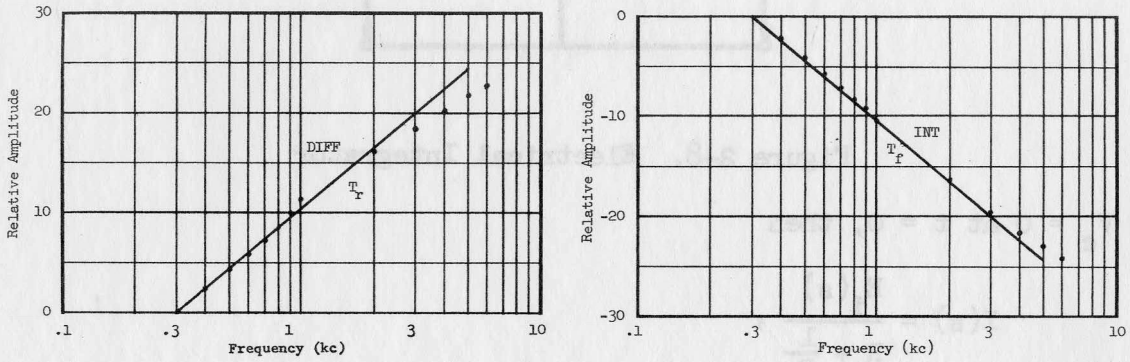


Figure 2-9. Amplitude Characteristics of the Integrator and Differentiator

The spectrum equalizer accentuates the low- and high- frequency components of a speech signal so that its power spectrum is constant.* This was accomplished by means of a double integrator (12 db per octave tilt) effective below 500 cps and differentiator active above 500 cps.

Symbolically the differentiator, integrator and equalizer will be denoted by T_r, T_f, and E, respectively. Degrees of these types of processes are not implied.

2.4.2.3. Automatic Volume Control

The amplitude compression device used in this investigation was an AVC employing variable attack levels and decay times. The attack time was of the order of 3 milliseconds. The circuit diagram and transfer characteristic are presented in Appendix D. The shape of the transfer

*Shyne (1960)

curve was a function of gain and threshold control, both of which were held constant for this investigation. Amplitude response was independent of decay time; the unit was useable for audio signals only. The symbol "V" will be used to denote the AVC operation.

This chapter has been devoted to definitions, derivations, and theoretical considerations. Various properties of speech signals and transmission systems were presented; a definition was made of the parameter λ, which is a measure of signal-to-noise ratio useful for peak-limited systems. The chapter was concluded by a discussion of the applicability of several signal processing devices.

3. EQUIPMENT AND EXPERIMENT DESIGN

"Human element" considerations dictate the necessity of experimental intelligibility evaluation. This is accomplished by presenting test materials to a trained group of "articulation test team" listeners. Preparation of test materials for the test team requires a representative system and special processing devices with parameters and functions scaled to enhance laboratory investigation. The purpose of this chapter is to outline the factors relating to equipment and articulation test design; a sample recording procedure is included.

3.1. System

For later explication of relatively complex recording details, the notation regarding SSB system components will follow that given in Figure 3-1 for the filter method. In the generalized SSB processing system, the blocks represent:

P^a - Audio processing (if any)

DG - DSB generator

P^d - DSB processing (if any)

F_m^s - SSB filter to derive SSB* signal

P^s - SSB processing (if any)

F_m^s - SSB filter to reduce processed SSB* signals to original bandwidth

N - Noise source

CH - Simulated channel

R - Receiver (frequency translation to audio spectrum space)

---

*In a strict sense of definition, these signals are not true SSB signals if processing is present at either D or S. In these cases, the "SSB" signals are not simple, frequency-translated versions of the input signal; hence, in cases where ambiguity might exist, the notation SSB will be used for these hybrid signals. Note that the foregoing remarks do not apply for P_a.

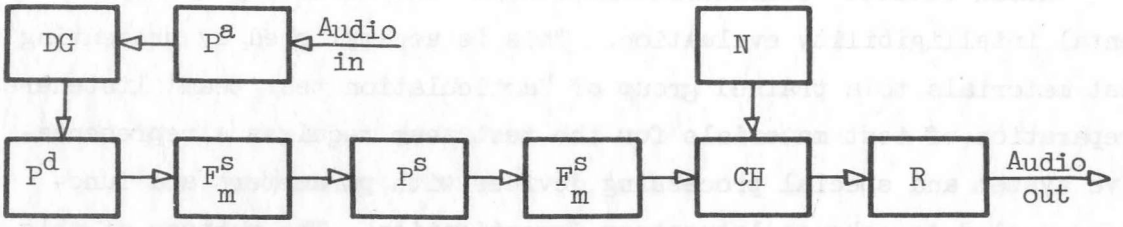
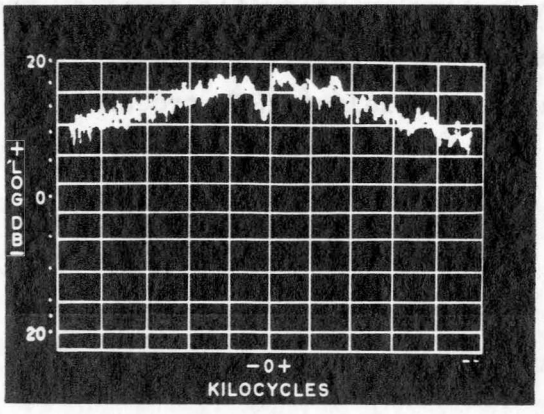
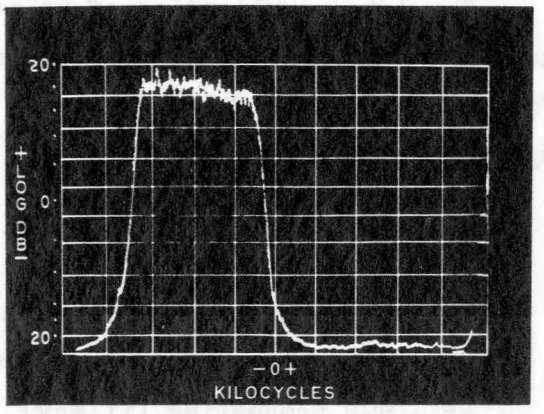


Figure 3-1. Filter Method of SSB Generation and Processing





a). b).

Figure 3-2. Spectral Studies of the SSB System (No Processing)
Vertical full scale: 40 db
Horizontal scale: 1 kc/division with zero corresponding to 250 kc.

a). DG output with white noise input
b). SSB spectrum after F^s_m of signal in a).

For simplicity, subsidiary equipment is not included in this block diagram. In this investigation the same frequency standard (buffered by cathode followers) was used for both frequency translations to negate the possibility of frequency translational error.*

The design of the representative SSB system is exemplified in Chapter 2 and in Appendix C. Additional information regarding actual system performance is given in Figure 3-2 with white noise as input to the DG of Figure 3-1. The spectrum of the DG output is given in Figure 3-2a. This study was made with a Panoramic Products Panalyzor. Note that the vertical scale is in db. Although it is not indicated in this figure, carrier suppression was of the order of 40 db. If the output of the DG of part a) is applied to a mechanical filter, the SSB signal of Figure 3-2b results. Note that the upper sideband frequency components have been attenuated more than 40 db. Horizontal scale for these figures is approximately 1 kc per division.

3.2. The Articulation Test

At this point it is important to realize that the evaluation of a representative system transmitting aural speech signals infers the decision-making process in the human mind regarding intelligibility. The purpose of this section is to point out the intricacies involved in the preparation, administration, and scoring of an articulation test.

3.2.1. Equipment

Intelligibility data was obtained by use of the articulation test facility (described by McCoy, 1958) installed in a room shrouded with drapery material to minimize reverberations and annoying external noises.

*Nickerson (1958) pointed out the effect of this parameter on the intelligibility of speech.

The facility included a tape transport and power amplifier for presentation of speech materials to the subjects. Each of the twelve listener positions was equipped with a volume control (which could be disabled from the master console for presentation of speech materials at a fixed volume level). The subjects listened to the speech materials with PDR-8 headphones using large foam earpads.* Ampex 350 and 601 tape recorders were used for preparation and presentation of pre-recorded test materials for the listener positions.

3.2.2. Test Materials

The results of a past investigation (Nickerson, Miller, and Shyne, 1960) indicated that three of the voice materials tested--Fairbanks, Navy Command Words and Phonetically Balanced Word Lists--resulted in similar intelligibility scores at specific noise levels. Therefore, selection of one of these voice materials for use in this study depended upon other factors. The Fairbanks Word Lists were selected primarily because the rate of presentation was twice that of the similar lists (i.e., the NCW and PB word lists were presented at a rate of one word every five seconds, while the Fairbanks words were presented at a rate of one word every two and one-half seconds). It appeared expedient in terms of time required and reduction of listener boredom to utilize the Fairbanks word lists.

The Fairbanks Rhyme Test was developed at the Speech Research Laboratory, University of Illinois. It is of the recognition type with stimulus words drawn from a vocabulary of 250 common monosyllables (fifty sets of five rhyming words each). Each test list of fifty words consists of one word from each of the rhyme sets and is identified by a list number. The 100 test lists used in this study were recorded by

*A frequency characteristic for the headsets is given in Appendix E-1.

Mr. R. Klumpp of the U. S. Navy Electronics Laboratory (NEL), San Diego, California. The stimulus words are presented two and one-half seconds apart preceded by the preface word "write" (not considered to be a part of the stimulus material). Stimulus materials are presented 8 db above normal speaker effort as afforded by continuous monitoring of a B&K power level tracer. (The B&K trace is similar to a VU meter response).

Since the λ adopted for this investigation was based upon signal peak value, the peak value of each test word was of primary concern. Measurement of the "peak list word" (PLW - the word or group of words with the largest peak value in that list) for the 100 lists of test material used in this investigation indicated less than 2.5 db variation with the exception of Lists 11 and 56. Variation of word peak-values for List 74 (chosen at random) was found to be in the order of $9\frac{1}{2}$ db. Figure 3-3 presents the audio (dashed lines) and SSB (solid lines) word peak-values for List 74 with the lowest word as reference. Both audio and SSB peak values are presented, since a given word will not have the same peak value after frequency translation from baseband to narrowband, even with no gain change. Figure 3-4 indicates the PLW values for each of the 100 lists of test words; the lowest word on each reel was normalized to 0 db. Since PLW values throughout the 100 lists differ only by approximately 2.5 db, an entire reel of test material can be calibrated for peak value by means of any list. This introduces an error less than or equal to 2.5 db for any specific list. With the assumption that a 2.5 db error would not appreciably effect the results of testing, it was possible to calibrate the PLW for a given condition with one list and use the following lists without recalibration to record tests at various λ conditions.

3.2.3. Test Team

The articulation test team was composed of married and unmarried women of college age who were found to have normal hearing. An all-

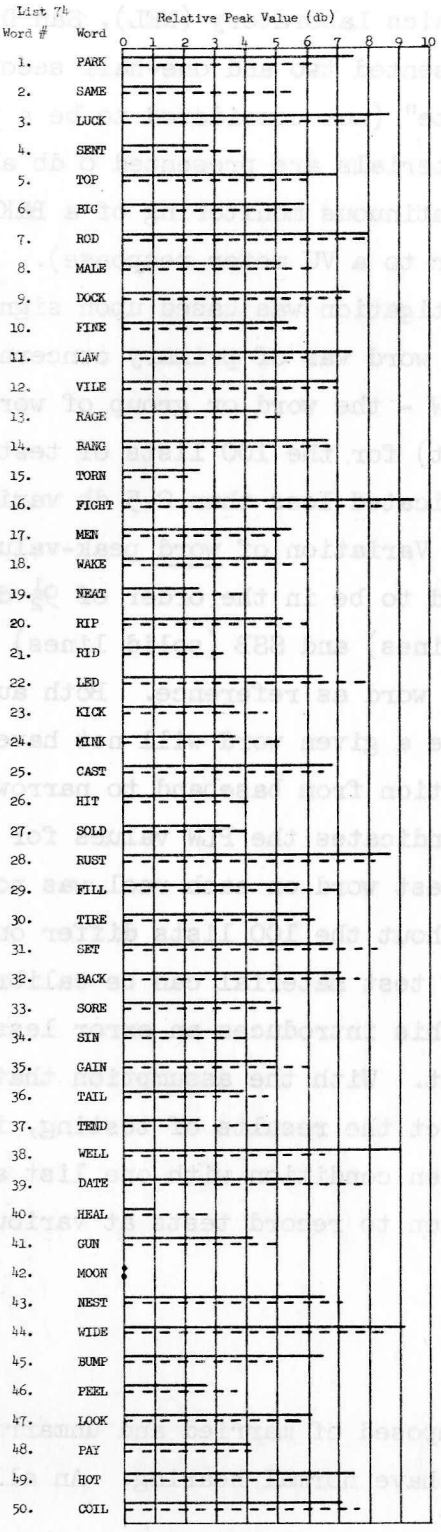


Figure 3-3. Relative Peak Values of Words of List 74.

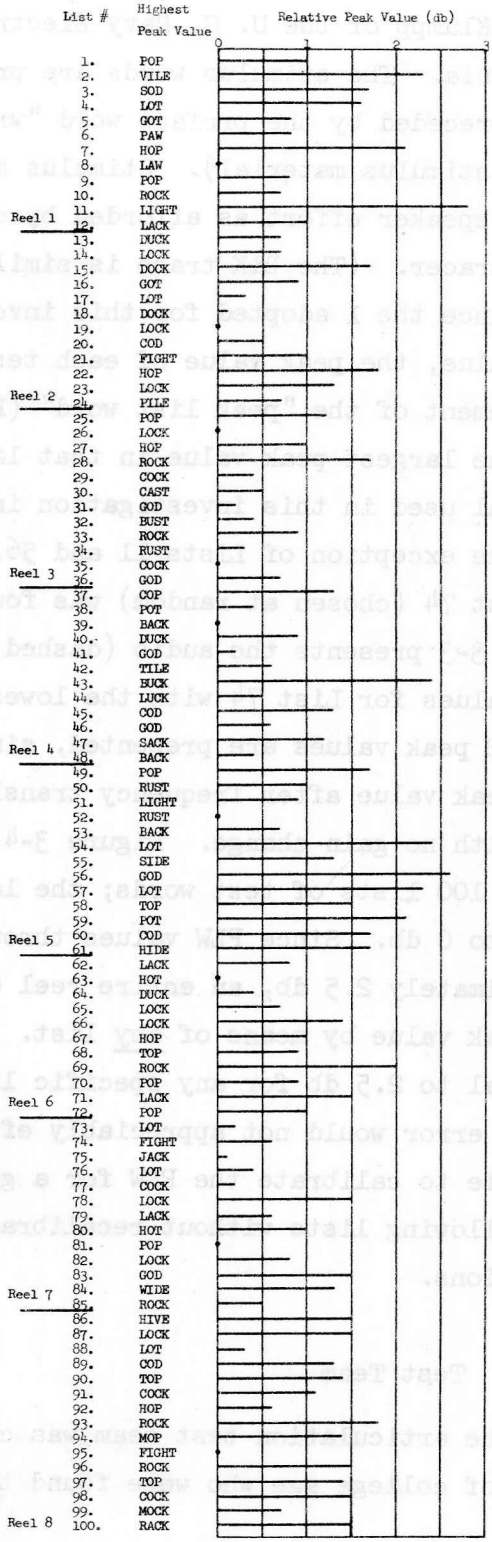


Figure 3-4. Peak List Words for Lists 1 - 100.

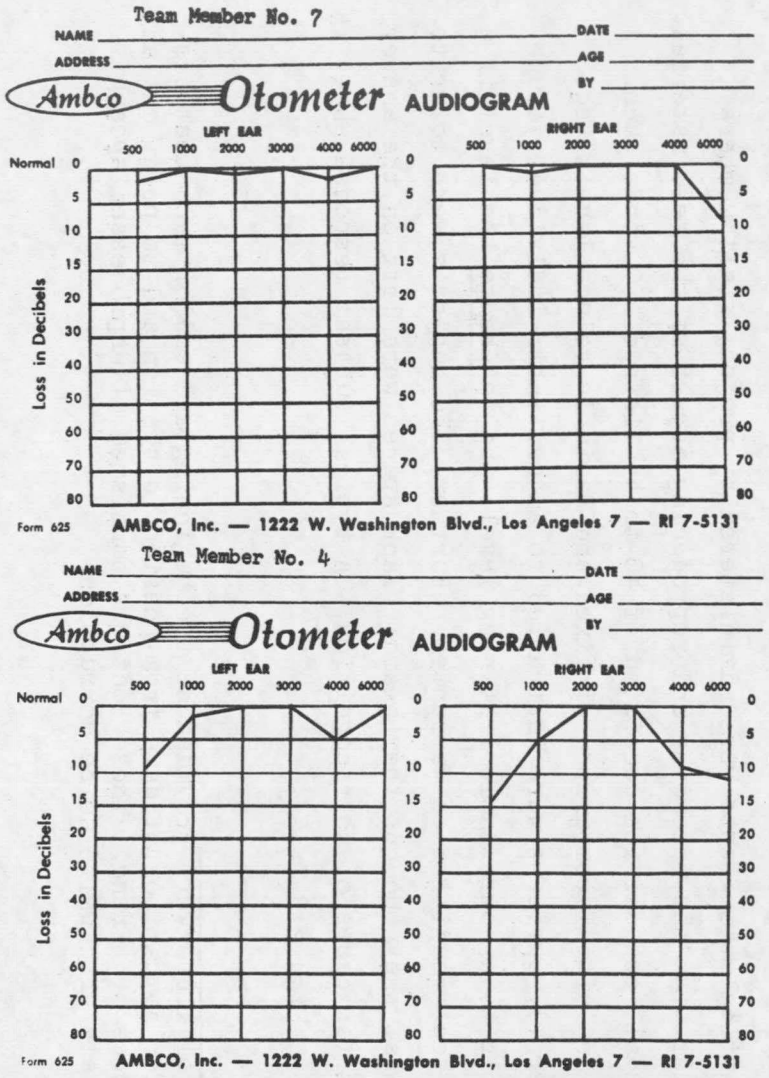


Figure 3-5. Audiograms of "typical" and "worst" Members Selected for the Test Team

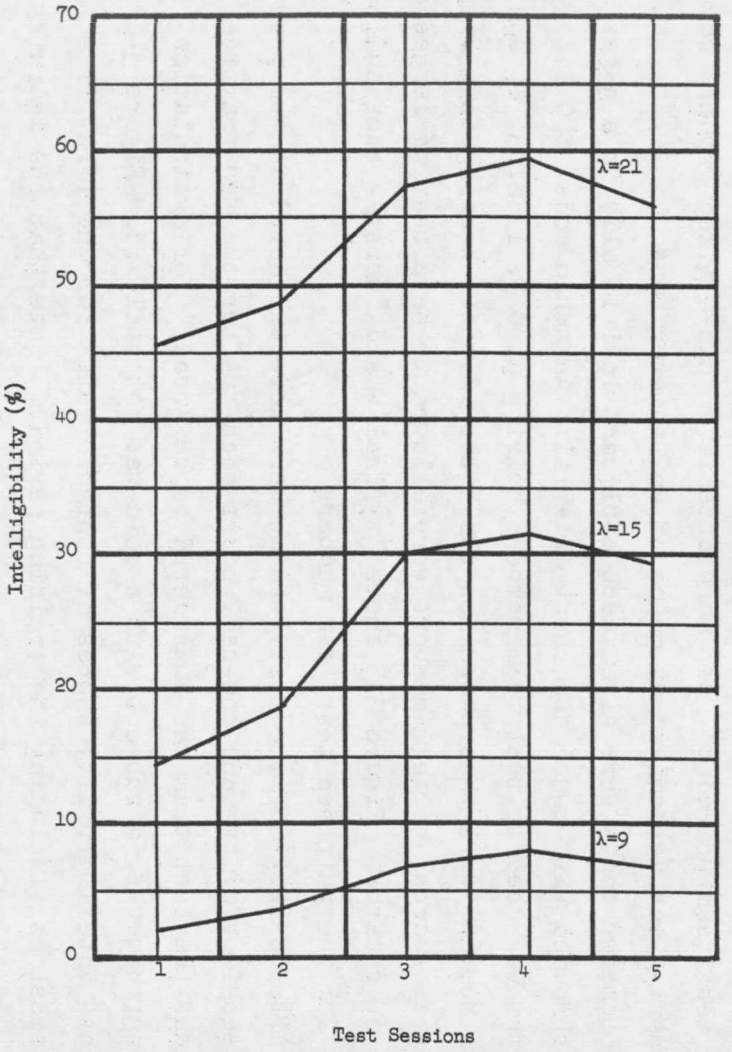


Figure 3-6. Control Scores for the First 5 Test Sessions for an Articulation Test Team

woman test team proved to be more attentive than a mixed group. Normal hearing was determined by means of an Otometer test.

Otometer tests were given each applicant individually in a quiet room shrouded with sound absorbent material. An Ambco model 600 D Otometer was used to test team members hearing at 0.5, 1, 2, 3, 4, and 6 kc. Members were then chosen on the basis of their hearing test with audiogram scores; no team members were allowed to have hearing losses of 15 db or greater. Figure 3-5 shows "typical"- and "worst"- case audiograms for the thirteen test team members.

After selection, the test team members were subjected to two practice sessions of two and one-half hours each. After the two practice sessions the test team was considered to be "ready" for articulation testing; however, a subject with a cold was not utilized. Figure 3-6 presents intelligibility scores of control lists* for the first five test sessions (including the training period). Note that the learning process was essentially complete by the end of the second training session.

Articulation tests were administered by project staff members familiar with the recording and reproducing aspect of the test material. A typical test session consisted of 40 to 44 lists of 50 words each. After an initial comparison of fixed and variable volume levels (see Chapter 4), subjects were instructed to adjust their individual volume controls to a comfortable listening level, and to change the level if a given test condition so warranted. Further, the subjects were to write down what they thought they heard; blank spaces were left on the answer sheets for cases of no recognition. A typical subject response sheet is shown in Figure 3-7.

*Control lists were provided at the beginning and end of each test session for comparison of data from session to session and to point out any evidences of tiring. These controls consisted of unprocessed speech materials with white noise corruption.

Test team answer sheets were scored by means of a separate group of four people who compared the subjects' conclusions to master word lists. The percent intelligibility for each test condition was the number of correct words divided by the total number of words times 100%. A typical test score sheet is shown in Figure 3-8*; individual listener scores (number correct out of 50 possible words in a list) are entered for the particular test conditions with mean values given in the right hand column.

An effort was made at all times to maintain the interest of the articulation test team. A reasonable hourly wage was paid for the time spent at test sessions. Test sessions were limited to approximately two and one-half hours with two ten-minute rest periods. Subjects were informed of the project goals and any new results. Control sheets with individual scores were kept from session to session so that a sense of competition was generated between members. Tours of the laboratory and the recording facilities were used to generate and maintain enthusiasm of the group.

3.3. Tape Preparation and Sample Articulation Test

All evaluations of processes in this investigation were performed using the λ definition of signal-to-noise ratio of (2-6):

$$\lambda = 20 \log_{10} \frac{E_p}{E_n} ,$$

*The first test condition, S_X, was an unscored control inserted in the latter testing phases to provide better testing stability. It allowed the subjects to "get adjusted" before the actual tests began; it was inserted after evidence indicated consistently low scores on the pre-control lists.

ANSWER SHEET FOR WORD TEST

MSC

Test No. 344, 86 Name 11 Date

1.	bin	26. jar	
2.	pale	27. wig	
3.	fit	28. pig	
4.	warn	29. kill	
5.	hop	30. sent	
6.	bread	31. wet	
7.	pink	32. fell	
8.	red	33. fit	
9.	gat	34. deal	
10.		35. fast	
11.		36. Cold	
12.	tuck	37. seal	
13.	rake	38. dark	
14.	line	39. for	
15.	save	40. feet	
16.	hop	41. pot	
17.	sia	42. foil	
18.	get	43. moon	
19.	save	44. feet	
20.	black	45. left	
21.	hide	46. file	
22.	med	47.	
23.	dock	48. coil	
24.	mone	49. pain	
25.	just	50. none	

Figure 3-7. A Typical Subject Response Sheet

CONDITION		SUBJECT INTELLIGIBILITY												SUMS AND MEANS		
		1	2	3	4	5	6	7	8	9	10	11	12	S	S*	M%
#344 CONTROL (S_A)				(PAW 5 CORES)												
LIST	λ^s															
86	15	13	12	18			22	16	12	14	15	13	13	148		
87	9	4	3	1			5	0	0	0	1	3	1	18		
#345 CONTROL ($N^P F_m^s$)																
LIST	λ^s															
88	15	14	12	19			14	18	15	17	16	16	13	154	311	31.1
89	9	5	3	4			6	2	6	1	7	2	1	37	81	8.1
90	21	22	26	36			29	32	28	32	28	28	21	282	567	56.7
#347 ($N^P D C_{N_2}^s F_m^s \lambda$)																
LIST	N_2 λ^s															
1	12 15	18	22	30			30	30	26	28	29	33	25	271	539	53.9
2	12 9	10	13	13			12	16	11	4	8	12	10	109	189	18.9
3	36 9	29	29	27			30	28	29	29	26	29	24	280	556	55.6
4	36 15	41	43	43			47	45	46	44	46	45	45	445	894	89.4
5	24 9	31	34	31			35	35	25	23	29	30	29	302	492	49.2
6	24 15	45	45	39			45	45	45	40	45	39	43	431	813	81.3
7	12 15	26	26	32			23	33	26	28	23	30	21	268	✓	✓
8	12 9	15	10	8			8	8	5	6	6	9	5	80	✓	✓
9	36 9	27	31	24			29	30	27	23	27	30	28	276	✓	✓
10	36 15	44	46	44			45	46	44	43	46	45	46	449	✓	✓
11	24 9	17	14	23			22	22	19	16	19	21	17	190	✓	✓
12	24 15	33	39	35			43	41	36	34	39	40	42	382	✓	✓

Figure 3-8. Typical Test Score Sheet

where E_p is the peak value of the given signal and E_n is the rms value of white noise measured by an average-reading, rms-calibrated VTVM (HP 400C). Such noise measurements are in error (from true rms) by approximately 1 db; this is verified in Appendix E-2. E_p was determined by use of a Tektronix 515A Oscilloscope; calibrated by means of the HP 400C VTVM such that full-scale deflection was obtained for the desired peak-to-peak value of the PLW of the calibration list. Note the difference between this definition of peak value and that defined as "that value which is exceeded 0.1% of the time". In all cases the definition of λ assumes equal bandwidths of signal and noise.

Consider a representative recording set-up, $\mathbb{P}\, C_n^s F_m^s$ (n db of SSB clipping, no other processing), as shown in block notation in Figure 3-9. The audio signal is reproduced by the Ampex 601 recorder from the NEL test tapes.* This signal is translated to narrow-band by means of

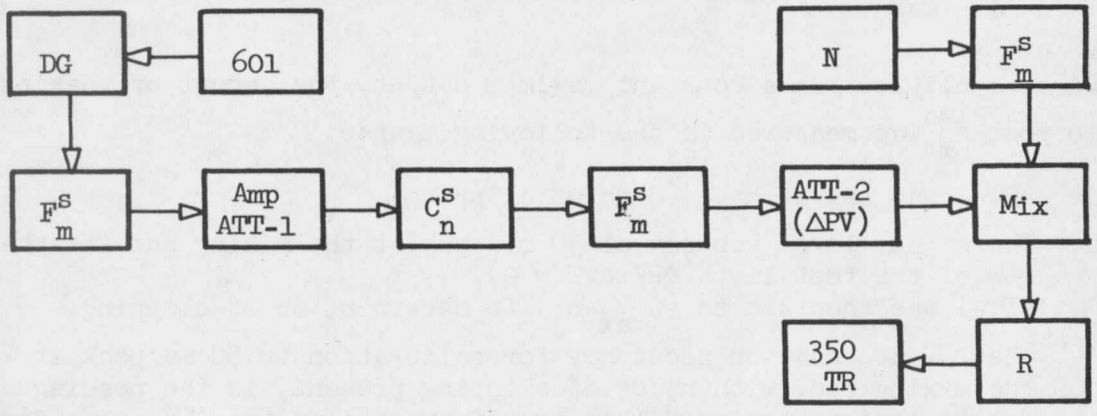


Figure 3-9. Representative Recording Set-up for n db of SSB Clipping, No other Processing.

*Test-tape λ was determined experimentally as approximately 45 db. Except for the λ^{27} conditions of Chapter 4, these "clean" test materials were not further corrupted prior to clipping.

the DG, and F_m^s removes the upper sideband. The SSB clipper is calibrated by means of the preceding amplifier and attenuator and the post F_m^s removes the higher-order frequency components produced by clipping. Recalibration after post-filtering is accomplished by use of the ΔPV attenuator. Filtered noise is introduced into the channel by the GR 1390-A Random Noise Generator and the filter, F_m^s. The resulting noisy signal is translated back to audio by means of the receiver. The receiver output is then recorded by the Ampex 350 for presentation to the articulation test team at a later time.

The SSB clipper was designed for clipping levels up to 24 db per unit with 50 mv peak input corresponding to 6 db of clipping. Calibration was accomplished at 50 mv peak so that the same scale on the HP 400C meter could be used for noise measurement and oscilloscope calibration. The amplifier at the input of the clipper was used to obtain 6 db of clipping with $(n_{max}-6)$ db of attenuation. Any desired clipping level $n_j (n_j < n_{max})$ could then be obtained by setting the attenuator at $(n_{max}-n_j)$.

Since a clipper has a constant maximum output, the amount of peaking due to post F_m^s was measured in the following manner:

1. ATT-1 was set at n_{max} (0 db of clipping)
2. The signal level was set at 50 mv peak at the mixing pad for the PLW of the test list (ΔPV ATT = 0).
3. ATT-1 was then set to $(n_{max}-n_j)$ to obtain n_j db of clipping.
4. The ΔPV attenuation necessary for calibration to 50 mv peak at the mixing pad, with n_j db of clipping present, is the peaking due to post filtering.[j] This is referred to as the ΔPV value for n_j db of clipping.

Since the amount of peaking due to post filtering is a function of translation and processing (as well as filter characteristics), ΔPV provides a valuable quality and accuracy check on recording procedures. Replication of identical processes normally produced ΔPV values with 1 db or less variation. The typical plot of ΔPV values for $\mathbb{P}C_n^s F_m^s$ is shown in Figure 3-10.

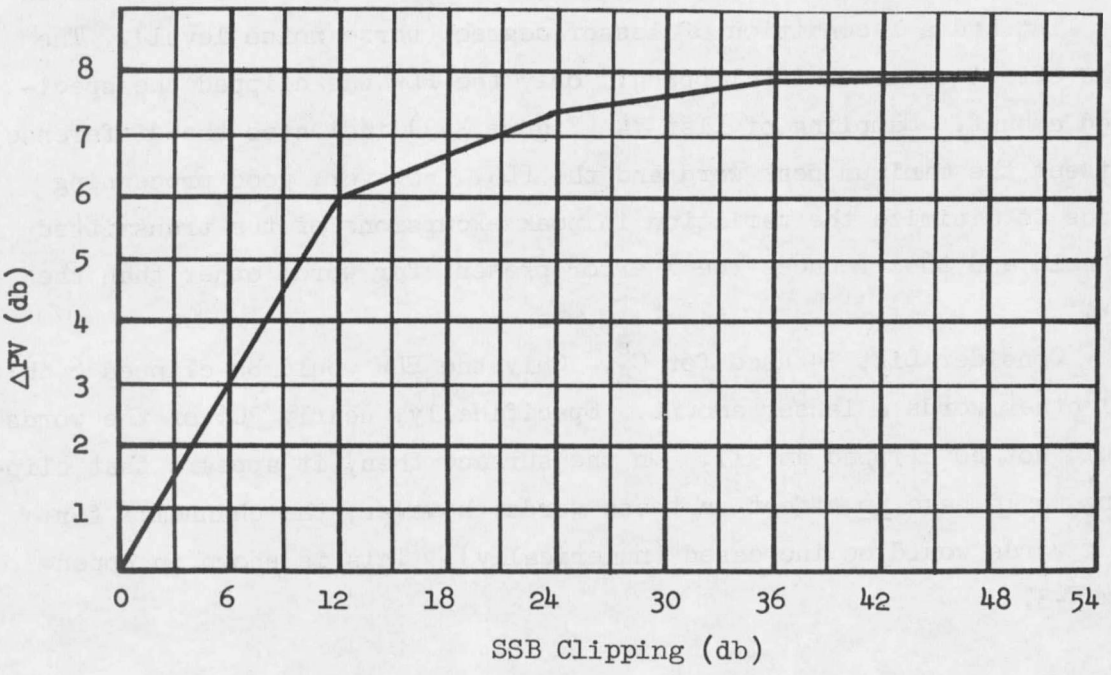


Figure 3-10. Peaking (ΔPV) Due to Filtering After SSB Clipping

The mixing pad was especially designed with internal signal and noise attenuators so that values of λ could readily be established. Calibration was accomplished with no attenuation of either signal or noise, so that full-scale oscilloscope deflection could be obtained and same-scale oscilloscope calibration and noise measurements could be made with the HP 400C meter. After this λ = 0 calibration ($\frac{E_p}{E_n} = 1$), the noise was attenuated 27 db to produce λ = 27. Attenuation of E_p in 6 db steps produced λ = 21, 15, 9, and 3.*

*Initial articulation tests were accomplished by holding signal constant and changing noise to effect divers λ. It was found, however, that the changing level of masking noise was perturbing to the subjects and required more sophisticated manipulations of volume controls from test to test. For this reason the signal level was altered (in 6 db steps) to change λ.

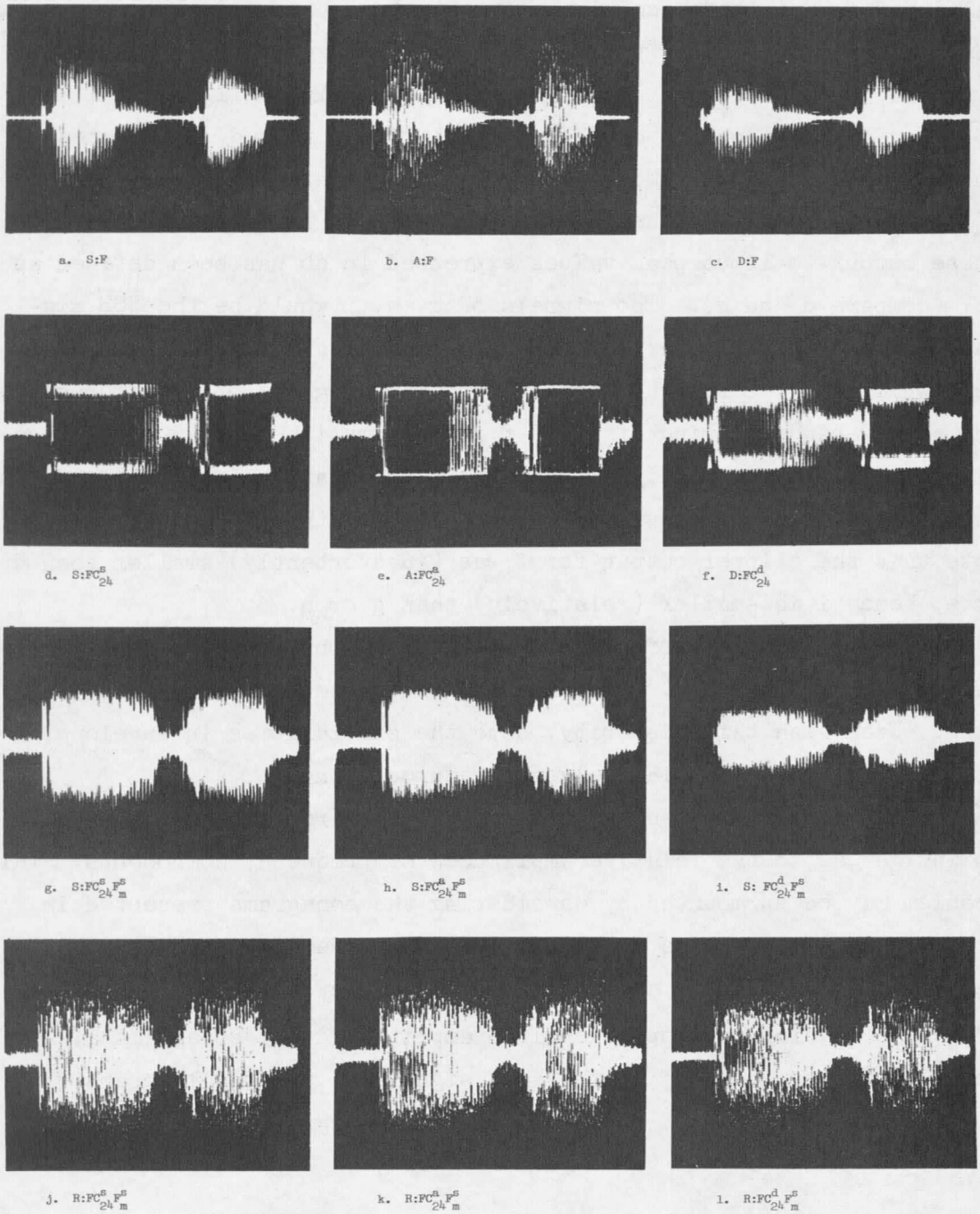
Since all calibration procedures used the PLW, all other words in the list had a λ condition of lesser degree (worse noise level). The same effect occurred in clipping: only the PLW was clipped the specified amount. Sampling of List 74 (Figure 3-3) indicates the difference between the minimum peak word and the PLW. However, good processing tends to minimize the variation in peak excursions of the transmitted signal, and thus reduces the λ error present for words other than the PLW.

Consider List 74 used for C_6^s. Only the PLW would be clipped 6 db; all other words a lesser amount. Specifically, nearly 30% of the words would not be clipped at all. On the surface then, it appears that clipping would have no effect on these words; however, the channel λ for such words would be increased (numerically). This is shown in Appendix E-3.

3.4. Signal Tableaux

Chapter 2 and the foregoing sections of this chapter have been devoted to the definition and explanation of the various properties and uses of signal processing as applied to a representative SSB communication system. At this point, perhaps some insight into the possible results of certain processes may be gained by the examination of actual signal samples, viewed in the representative system. We have already seen the time domain graphic of an audio signal in Figure 2-1, but what is the form of a narrow-band version of an audio signal? Consider Figure 3-11a. This is an oscillogram of the word "poem-fate"* as an SSB signal (the "poem-fate" audio input signal is shown in Figure 3-11b; it is bandlimited 300-3000 cps, F). In Figure 3-1 this signal would appear at the output of the first mechanical filter if $p^a = p^d = 0$.

*The word "poem-fate" was chosen for content: initial plosive, long vowel sounds, glide, fricative, and final plosive.



a. S:F b. A:F c. D:F

d. $S:FC^{s}_{24}$ e. $A:FC^{a}_{24}$ f. $D:FC^{d}_{24}$

g. $S:FC^{s}_{24}F^{s}_{m}$ h. $S:FC^{a}_{24}F^{s}_{m}$ i. $S:FC^{d}_{24}F^{s}_{m}$

j. $R:FC^{s}_{24}F^{s}_{m}$ k. $R:FC^{a}_{24}F^{s}_{m}$ l. $R:FC^{d}_{24}F^{s}_{m}$

Figure 3-11. Oscillograms of the word "poem-fate" for various system samples and processes. These oscillograms are to the same horizontal and vertical scales, and the SSB filters of Figure 3-9 were adjusted for unity gain. Note that no direct correlation may be made between the baseband signals (b,e,j,k,l) and the narrow-band signals (a,c,d,f,g,h,i,), since frequency translation (and probably gain) was interposed.

Note that in these oscillograms the units of the vertical and horizontal axes are amplitude and time respectively. Figure 3-11c depicts the time-domain function of "poem-fate" as a DSB signal. Figures 3-11 d, e, f are clipped versions of a, b, c respectively, while g, h, i are $\overline{\text{SSB}}$ viewed at the output of a mechanical filter with d, translated e, and f as filter input. Notice that with unity sinusoidal gain in the F_m^s, the output has a higher peak value than the input. The ratio of these output-to-input peak values expressed in db has been defined as ΔPV (compare d and g). The signals of g, h, i would be the $\overline{\text{SSB}}$ signals applied to the power amplifier for transmission in an actual system, while j, k, l are the respective output signals of an SSB receiver. The narrow-band clipping of d and f was accomplished at 250 kc; these signals are not sharp clipper outputs, since peaking was introduced in the amplifier (4 mc bandwidth) used to amplify the clipped signals. Note that the clipper output for f was (inadvertently) smaller than d or e; hence i is smaller (relatively) than g or h.

The oscillograms of b, e, j, k, l are of broadband signals (baseband), while a, c, d, f, g, h, i are oscillograms of narrow-band signals. Even with this disparity, note the similarities in envelope, even in the cases of relatively severe processing.

In a sense, the oscillograms of 3-11 are restrictive; they give no insight as to the relative amplitudes of frequency components. This problem may be surmounted by considering the sonagrams presented in Figure 3-12.* Sonagrams may be considered three dimensional: the vertical and horizontal axes represent frequency and time respectively, while shading is proportional to amplitude. The sonagrams a, b, and c of Figure 3-12 depict the same signals as the oscillograms of b, j and k respectively, of Figure 3-11. The sonagram of d depicts the

*The sonagrams were prepared by Mr. R. A. Houde, through the facilities of the Acoustics and Life Sciences Dept., Research Division, General Dynamics | Electronics.

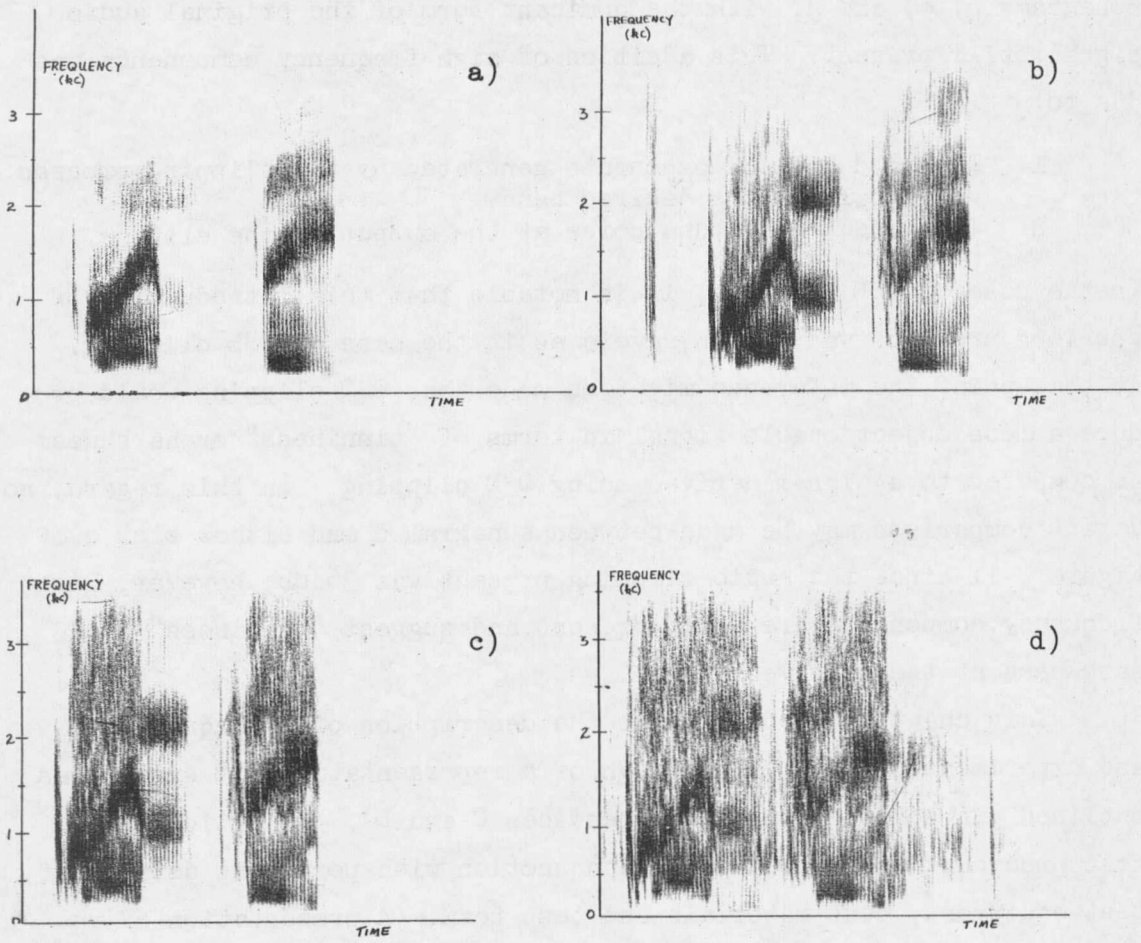


Figure 3-12. Sonagrams of "poem-fate" for:

a) An audio signal (NP)

b) Receiver output with C_{24}^{s}

c) Receiver output with C_{24}^{d}

d) Receiver output with C_{36}^{a}

receiver output with 36 db of audio clipping (no other processing). Note that additional high-frequency components were introduced in the sonagrams b, c, and d, with the dominant form of the original audio signal still present. This addition of high frequency components was due to:

1. Intermodulation components generated by the clipping process which fall in the desired band.
2. Increased background noise at the output of the clipper.

In the case of SSB clipping, it is notable that this introduction of spurious products was not as severe as in the case of DSB clipping. At this point the inference might be made that SSB clipping would produce a less objectionable signal in terms of "tinniness" or harshness as compared to a signal derived using DSB clipping. In this regard, no direct comparison may be made between sonagram d and either b or c of Figure 3-11 since the audio clipping present was 36 db; however, high frequency components are again present and suggest "tinniness" and harshness of the received signal.

This chapter was devoted to the description of the equipment and experiment design. The design of a representative SSB system was outlined (in conjunction with Appendices C and D). An articulation test procedure was presented in conjunction with pertinent details of test equipment, test materials and test team. A presentation of recording details and precautions, a sample articulation test and a graphical signal analysis completed the section. We are now prepared to consider the experimental results of the study.

4. DATA PRESENTATION AND INTERPRETATION

In the preceding chapters, the foundations have been laid for the consideration of signal processing techniques as applied to SSB systems. This chapter will be devoted to the presentation and analysis of the experimental results which were obtained.

4.1. Data Reduction Sample

The data pertaining to the evaluation of a particular type of signal processing as obtained from the articulation test team were in numeric form. The number of correct responses of one subject for a list of fifty words represents one sample score (N = 1). Sample scores (doubled to obtain percentage intelligibility) for two different test conditions are:

1. $\mathbb{P}C^{s}_{24}F^{s}_{m}$, $\lambda^{s} = 9$:

   I = 56,56,62,54,64,62,54,58,50,56,60,66,60,50,62,80,56,
   54,60,76,44,38,72,32,42,54,40,58,52,40,38,36,54,28,
   34,38,42,52,32,32,46,44,70,26,34,44,44,46,34,54,42,
   66,46,52,42,52,62,52,40. (N = 59).

2. $\mathbb{P}C^{s}_{12}F^{s}_{m}$, $\lambda^{s} = 3$:

   I = 2,2,0,4,2,0,0,0,0,2,4,6,6,2,0,0,2,2,0,10,2,2,4,2,2,
   + 34 0's. (N = 59).

These scores are plotted as frequency polygons in Figures 4-1 and 4-2. Note that the distribution of scores near the center of the intelligibility range tends to approach the normal distribution, while at an extremal point they do not (this extremal "bunching" of points was also evidenced as intelligibility scores approached 100 per cent).

Considering such data, derived from an unknown population, it was necessary to make a choice of a measure of central tendency. The desirable characteristics of such a measure were assumed to be:

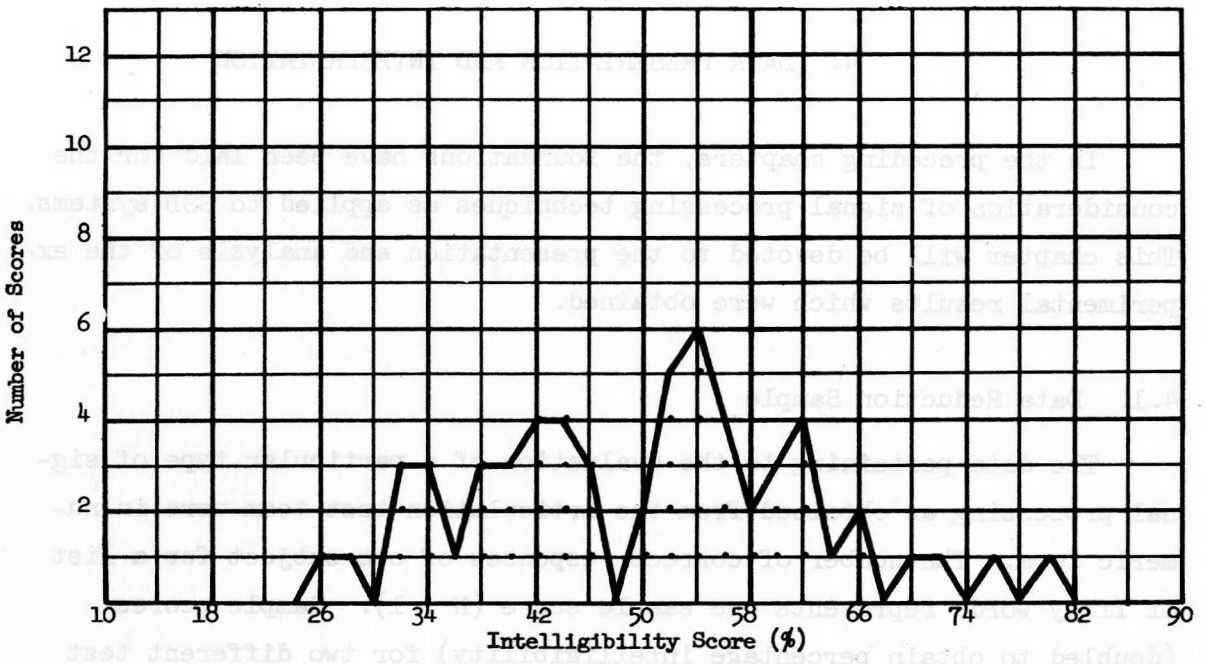


Figure 4-1. Frequency Polygon for $\mathbb{P}C^{s}_{24}F^{s}_{m}$, $\lambda^{s} = 9$

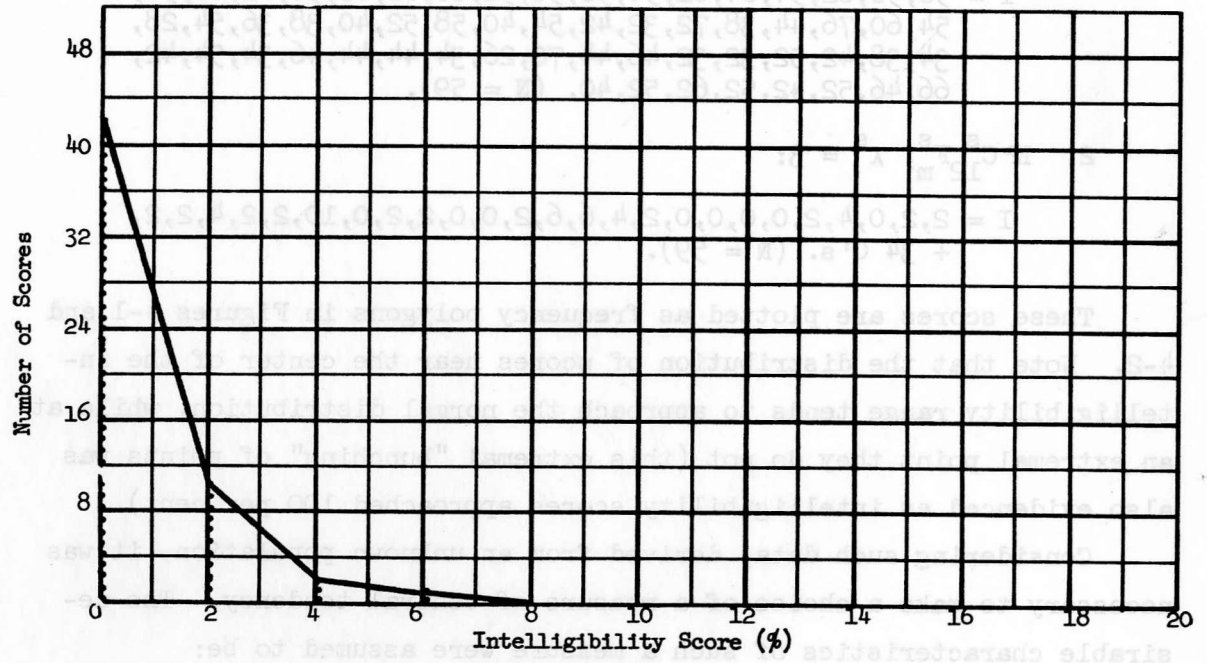


Figure 4-2. Frequency Polygon for $\mathbb{P}\ C^{s}_{12}F^{s}_{m}$, $\lambda^{s} = 3$

1. Sampling stability
2. The use of all sample scores
3. Perspicuity
4. Simplicity of computation

Using these characteristics as criteria, it was decided to use the mean as a measure of central position. In this way relative standing could be indicated in terms of the mean and this could be done without reference to the distribution of original scores. The mean is defined as:

$$\bar{x} = \frac{\Sigma x_i}{N},$$

where N is the total number of sample values and the x_i are the individual values of these intelligibility samples. The means for the above examples are 50.0, 1.0, respectively; N = 59 for each case. For purposes of comparison, the mode and median of these examples are 54, 52; and 0, 0.5, respectively.

Later, it will be of interest to determine whether or not a significant difference exists between two given points on a graph. Since it was impossible to assume a particular distribution for any given sample, it was necessary to employ a nonparametric test of significance when comparing the sample distributions represented by any two points of a graph. Since at least ordinal measurement was achieved in the data collection for this experiment, the Mann-Whitney U test (as outlined by Siegel, 1956) was used to determine whether two independent groups of sample scores (for different parametric conditions) had been drawn from the same population. This test is applied specifically to sample points of Figure 4-13; the method and computational details are given in Appendix G.

4.2. Test Results

The intelligibility scores for various system processes will be presented in essentially the same chronological order as they were

derived. The study was divided into four basic parts:

1. Preliminary audio processing
2. Audio (premodulation) processing for SSB
3. Narrow-band processing for SSB
4. Iterative, interdependent and other combinational processes

If possible, an interpretation will be made as each set of data is presented; tabular data for N and $\bar{x}$ for each graphical presentation is given in Appendix F.*

4.2.1. Preliminary Study

The early efforts of the study were concerned with:

1. Audio processing (no SSB system was used)
2. 5 kc bandwidth signals
3. λ determination at audio (λ^a)

All discussions in this section will be based on these restrictions.

The first question that arose concerned the required values of λ to yield control intelligibility scores in a median range (say, 10 - 90% or 20 - 80%). Figure 4-3 shows the scores obtained on the first trial. (Note that these intelligibility scores are for single words as opposed to scores for continuous speech). The scores for $\lambda = 0$ were consistently zero, and the curve ranged upward and to the right as shown. This did not adequately depict the higher intelligibility levels, so $\lambda = 9, 15, 21, 27$ was used on the second trial. The results are shown in Figure 4-4.

An attendant problem in this original determination was the absence of criteria to decide whether the test materials should be presented to the test team members at a fixed volume level, or whether the team should be given the freedom to choose a comfortable listening

*For convenience the figure numbers of this chapter will correspond directly to the table numbers of Appendix F. For this reason Figures 4-1 and 4-2 are assigned the trivial Tables F-1 and F-2.

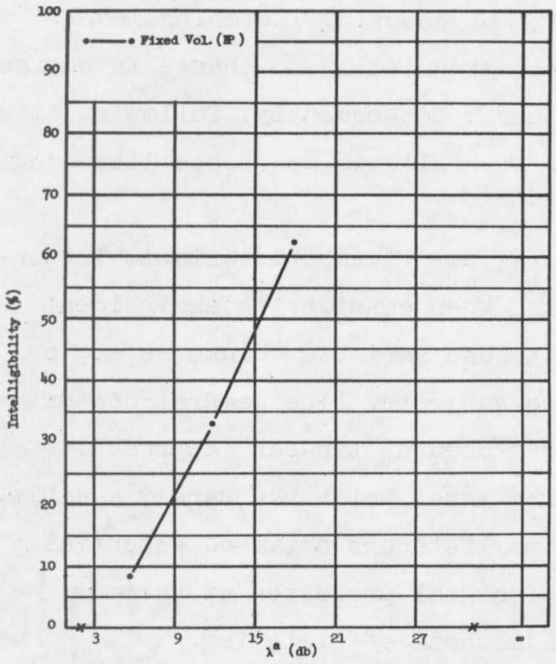


Figure 4-3. Intelligibility Data for Initial Control, λ^a.

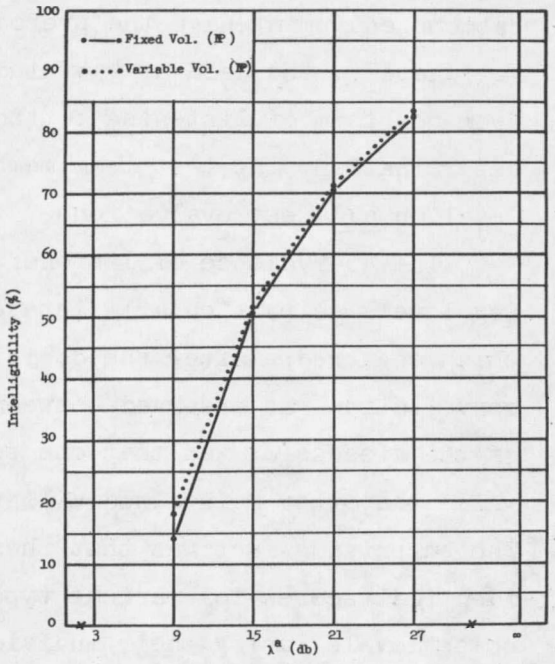


Figure 4-4. Modified Control - λ^a Presentation

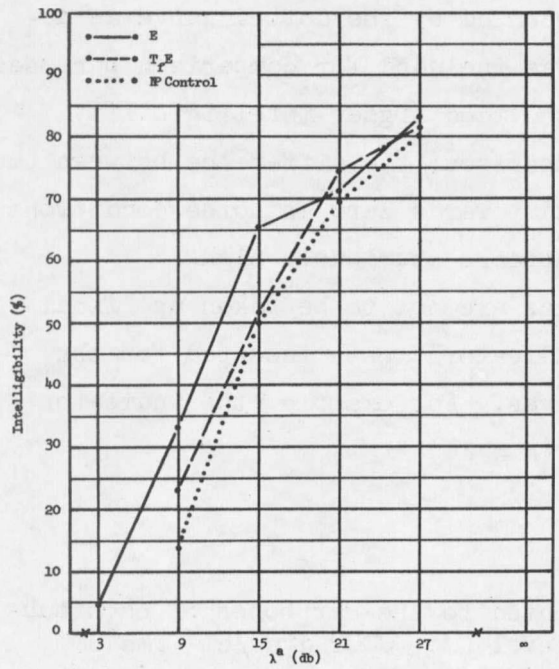


Figure 4-5. E and $T_F E$ (λ^a)

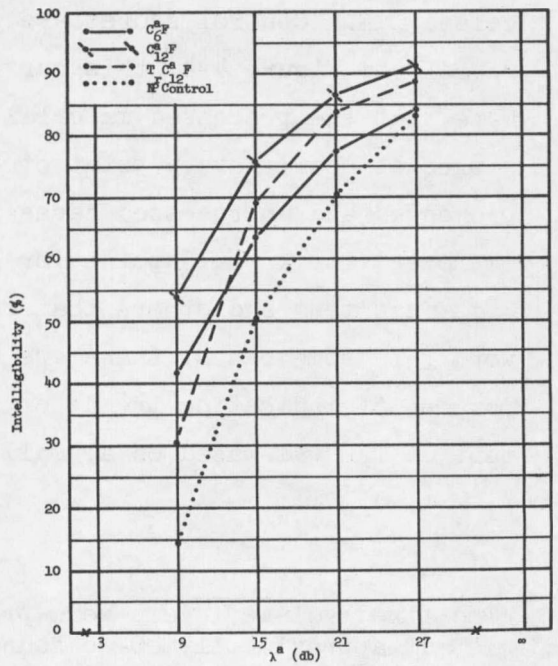


Figure 4-6. Sample Study of Audio Clipping, (λ^a).

level. The latter seemed reasonable since operators of most communications equipment have the prerogative of choosing listening level. Fortunately, the data of 4-4* indicate essentially no change in scores from one form of listening to the other. Consequently, following this determination, the test team members were allowed to choose listening level on all test evaluations.

The equivalence of listener scores for fixed and variable listening levels is particularly interesting when considering an incident which occurred: after the decision to use variable volume level, a conversation was observed between two subjects. The general context of the discussion was that one subject used a "tin-ear" approach, while the other maintained a listening level which was barely audible. The surprising fact was that these two listeners obtained essentially identical scores for various types of signal processes at various noise levels, using their individual methods of listening.

Figure 4-5 presents data of intelligibility as a function of λ^a for the equalization process (E) and the combination of linear processes $T_f E$. Control scores are presented by the dotted curve as derived from Figure 4-4; this curve was included for comparison purposes. Note that the processed materials provided higher intelligibility scores at nearly every point of comparison; the difference between the processed and unprocessed cases tend towards zero as noise conditions improve. The $\lambda^a = 15$ point for E appears spuriously high.

This test and others that follow are not to be taken as "final word"; in some cases, these are "indicator" tests inserted for the purpose of indicating trends or ranges. For example, the indicator test of $T_f E$ was based on a small N (N = 10).**

*The fixed volume levels were presented to the earphones of each subject at approximately 80-db sound level (re .0002 dyne/cm^2 rms).
**When N is small, the variability of mean scores can vary significantly from member-to-member of the test team and from test-to-test.

A sample study of audio clipping is given in Figure 4-6; two degrees of clipping, T_r prior to clipping and control are presented. The indications seem to be:

1. C^a_{12} appreciably increases intelligibility scores at all λ, and in the 50% to 80% intelligibility range provides 6 db or more increase in signal power over unprocessed speech.
2. For good listening conditions, $T_r C^a_{12}$ approaches C^a_{12}.*
3. At $\lambda = 9$ the indication is that C^a_6 produces higher scores than $T_r C^a_{12}$; this might lead one to suspect that the $\lambda = 9$ point for C^a_6 is higher than it actually should be.

The question might be posed: For an AVC circuit, do optimum values of decay-time and amount of suppression exist? In order to obtain estimates to at least partially answer this question, the data of Figures 4-7 and 4-8 were obtained. First, with a constant attack level equal to 0.5 volts p-p (which may be referenced to the amount of suppression by means of Figure D-7) intelligibility scores were obtained for various values of decay-time. These data are presented in Figure 4-7. The intelligibility scores tend toward a maximum in the left portion of these curves. So, somewhat arbitrarily, the decay-time for optimum performance was chosen to be 0.5 seconds for all λ levels. Then, holding this value of decay time, the data of Figure 4-8 were obtained by changing the input signal peak-value from test-to-test. Scores for the various λ levels indicate that the choice of 0.5 volts peak-to-peak or Figure 4-7 was judicious.**

*Contrary to popular belief, this test (and others which follow) indicate that differentiation (T_r) prior to clipping does not increase intelligibility scores.

**Actually an indicative test matrix was run about this point.

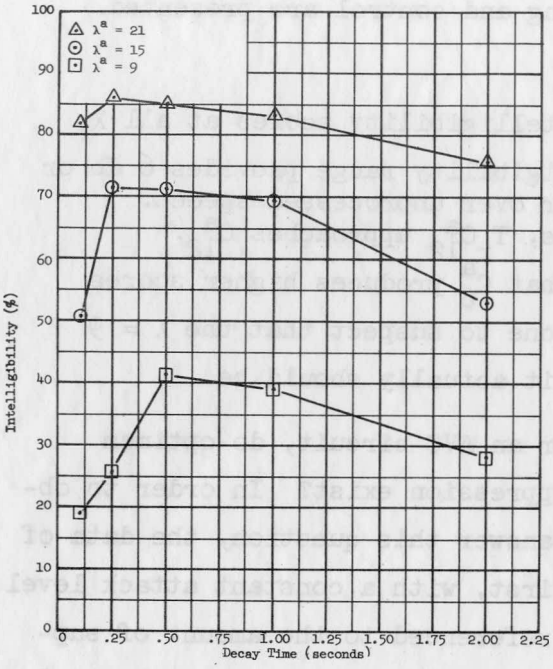


Figure 4-7. Intelligibility Test of Audio AVC for Maximum Input Voltage of 0.5 v peak-to-peak (λ^a).

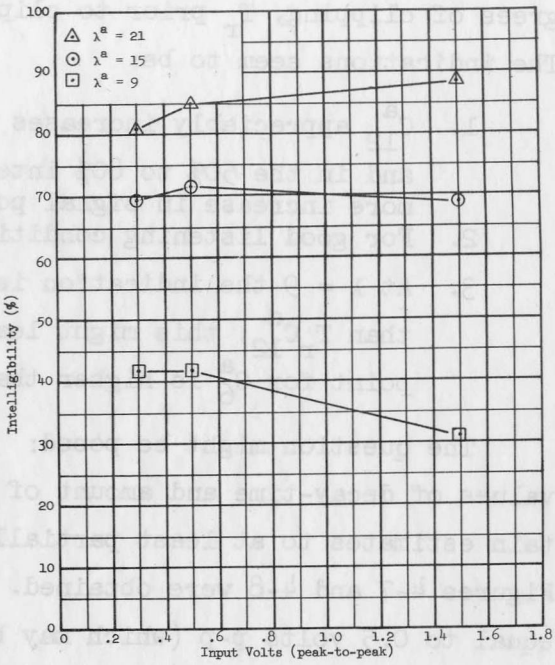


Figure 4-8. Audio AVC for Constant Decay-Time of 0.5 Seconds (λ^a).

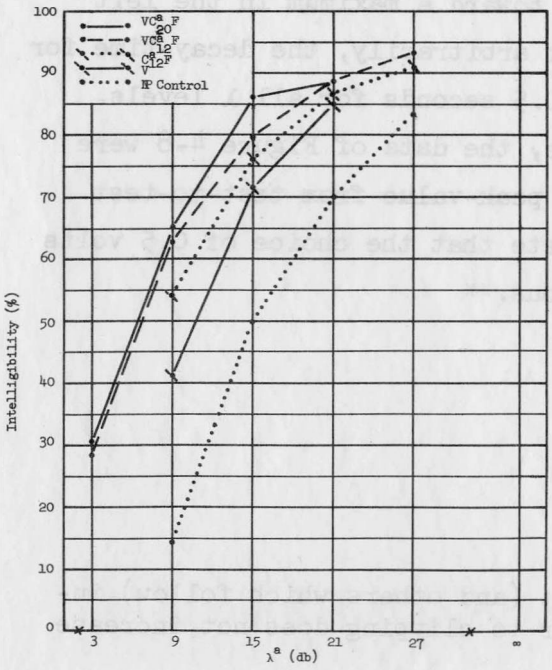


Figure 4-9. Combinations of V and C^a_n, (λ^a).

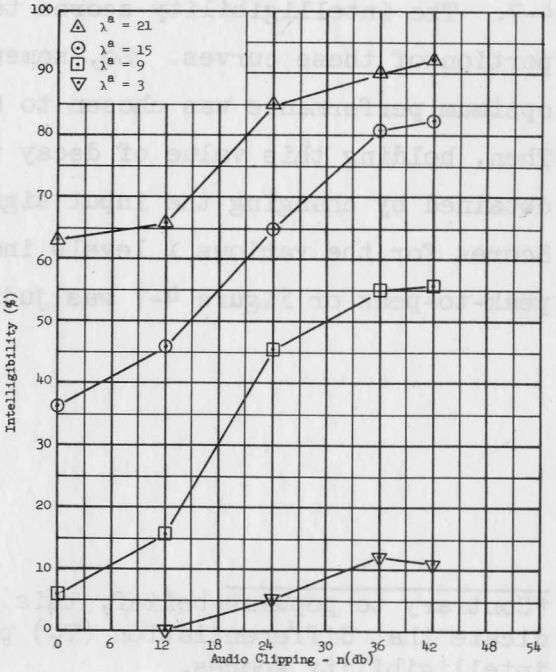


Figure 4-10. Intelligibility Estimates for $C^a_n F$ 3 kc Bandwidth (λ^a).

As a result of these data, the AVC parameters were fixed at 0.5 volts attack level and 0.5 seconds decay-time. The notation for the device in this condition will be V.*

Intelligibility as a function of λ for the V process is given in Figure 4-9. Also included in this figure are combinations of V and C^a as indicated (taken from Figures 4-6 and 4-7). Again, these are indicator tests and N is quite small (for $VC^a_{12}F$, N = 24; for $VC^a_{20}F$, N = 8). However, the indication is that in the vicinity of the 70% intelligibility the AVC is nearly as effective as 12 db of clipping; in this range and up to approximately 85% intelligibility, the combination of clipping and AVC affords higher scores than either process alone. An increased clipping level appears to further increase intelligibility scores.

4.2.2. Audio Processes for SSB

The SSB system utilized for this investigation was of 3 kc bandwidth. For purposes of comparison to the clipping data presented in the preceding section and outlining the general implications of a narrower bandwidth, data for C^a_n were obtained for 3 kc signals and are presented in Figure 4-10. Note that λ^a is the parameter in this case, and that, as in the preceding clipping case, intelligibility scores tend to increase as either n or λ^a increase. The parametric curves end at the n = 0 line; scores at these points are control (NP) scores since processing is zero. It appears that the points at n = 12 are low.

Aside from the above possible errors, a comparison of the n = 12 points of Figures 4-10 and 4-6 indicates a considerable disparity of scores. For example at λ^a = 15, 5 kc bandwidth yields an intelligibility score of 75.7% (I = 75.7) and 3 kc bandwidth yields I = 46.1.

*Note that for the determination of AVC parameters, the intelligibility scores were derived using single-word stimulus materials; it is probable (perhaps certain) that these parameters would not be the same for continuous speech.

Part of this difference may be attributed directly to a relative change in λ of 2.4 db* due to bandwidth reduction. Since the rms value of the noise (E_n) must be held constant, bandwidth reduction tends to make the noise condition more severe. Further loss of intelligibility must be expected since all frequency components in the 3000-5000 range have been removed.

Now consider audio processes as applicable in an SSB system. Historically, this transition caused problems. Probably the most important was the consideration of <u>where</u> λ should be defined for the system. In retrospect this seems like a trivial problem; however, a number of experimental evaluations were made with λ defined at the receiver output. This seemed reasonable, since all the preceding tests were concerned with baseband signals. No particularly striking trends existed; and the data collected is of little value unless audio processing (say for a public address system) is to be effected by using an SSB system (it works). Actually, it was assumed for these tests that λ was invariant with linear frequency translation; but, as indicated in Chapter II, this is not true. The data for λ^r will not be presented.

In an actual system, λ is determined in the channel; the laboratory system included a channel (although it was merely a piece of coaxial cable). Consequently, to adequately represent the actual state of affairs, λ was determined immediately following the last mechanical filter in the processing system. In order to represent selectivity of the receiver, the additive noise was band-limited with a filter having characteristics essentially identical to the last channel filter.** λ^s is thus defined.

With these considerations resolved, it is possible to proceed to consider Figure 4-11; this is one of the more important results of

*See Appendix E-4 for justification.
**See Appendix E-5.

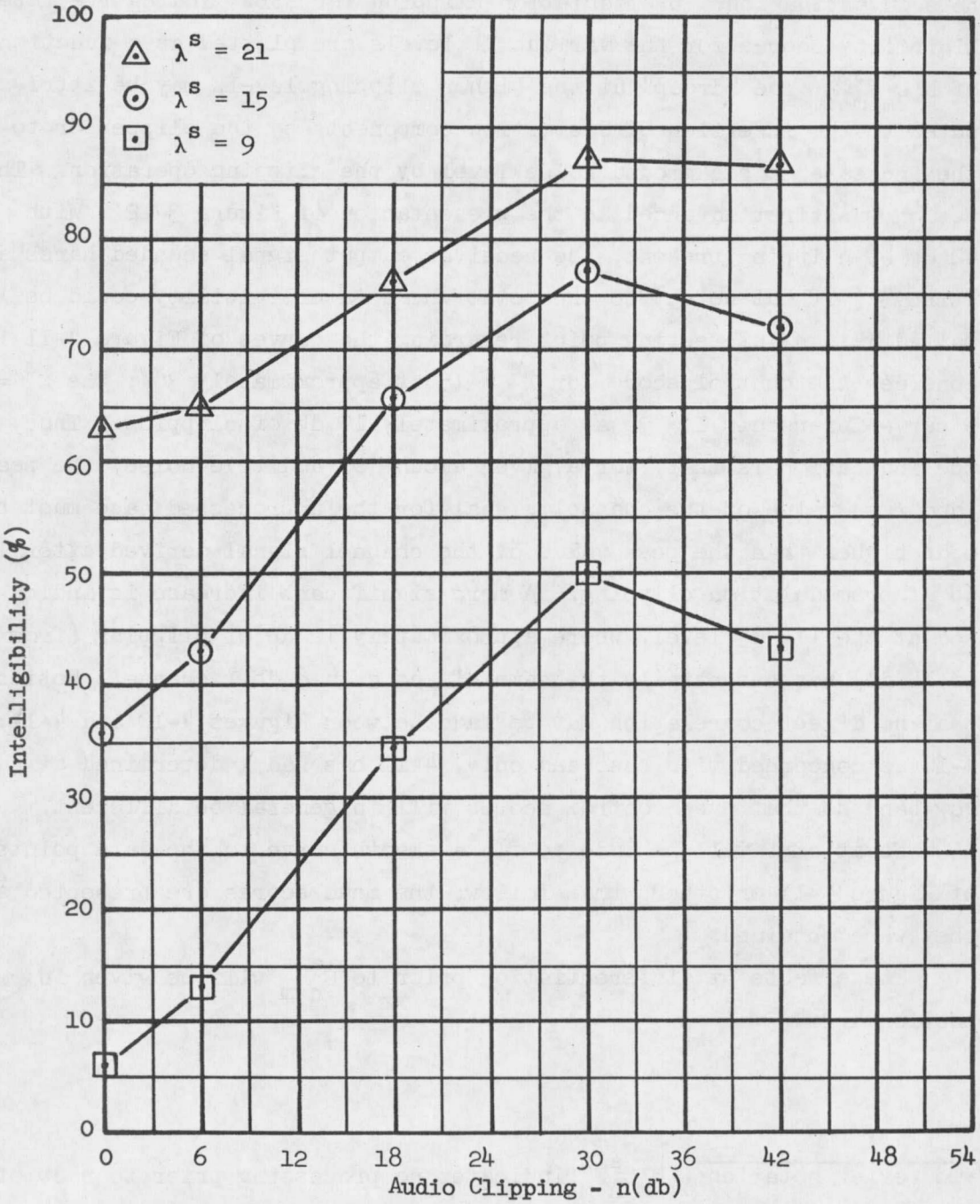


Figure 4-11. Intelligibility Estimates for $\mathbb{P}\, C_n^{a} F_m^{s}\ (\lambda^s)$.

this investigation: premodulation clipping for SSB. Indicative intelligibility scores for the various λ^s levels are plotted as a function of $\text{NP}\,C_n^a F_m^s$.* The "droop" at the higher clipping levels may be attributed to the insertion of distortion components by the clipper or to the increase of background noise level by the clipping operation. This notion was first advanced in the presentation of Figure 3-12. With 42 db of clipping present, the receiver output signal sounded harsh and "tinny" (without noise) to the point where intelligibility could be impaired.** An interesting point regarding the curves of Figure 4-11 is: consider the control score for $\lambda^s = 15$ at approximately 36%; the $\lambda^s = 9$ curve intersects I = 36 at approximately 19 db of clipping. Thus, the indication is that, for a given amount of additive noise, the peak (and rms) value of the channel signal for the unprocessed case must be 6 db higher than the peak value of the channel signal derived after 19 db of premodulation clipping. A more significant increase is indicated at the I = 65 level, where approximately 12 db of clipping (from 6 to 18 db) has essentially the same effect as a 6 db λ change. Notice that no direct correlation may be made between Figures 4-10 and 4-11; 4-10 is concerned with baseband only, 4-11 has had λ determined at narrow-band so that even control scores will in general be different.

No attempt will be made to fit a smooth curve to the data points of Figure 4-11 or others that follow; the mean scores are presented as they were obtained.

The effects of differentiation prior to $C_n^a F_m^s$ will be given in section 4.2.4.3.

---

*To review notation: $\text{NP}\,C_n^a F_m^s$ indicates no processing prior to n db of audio clipping; bandwidth is restored at narrow-band using an SSB filter.

**No quantitative tests were run on the "quality" of the speech signals after processing. However, intuitive opinions were formed by the project staff members.

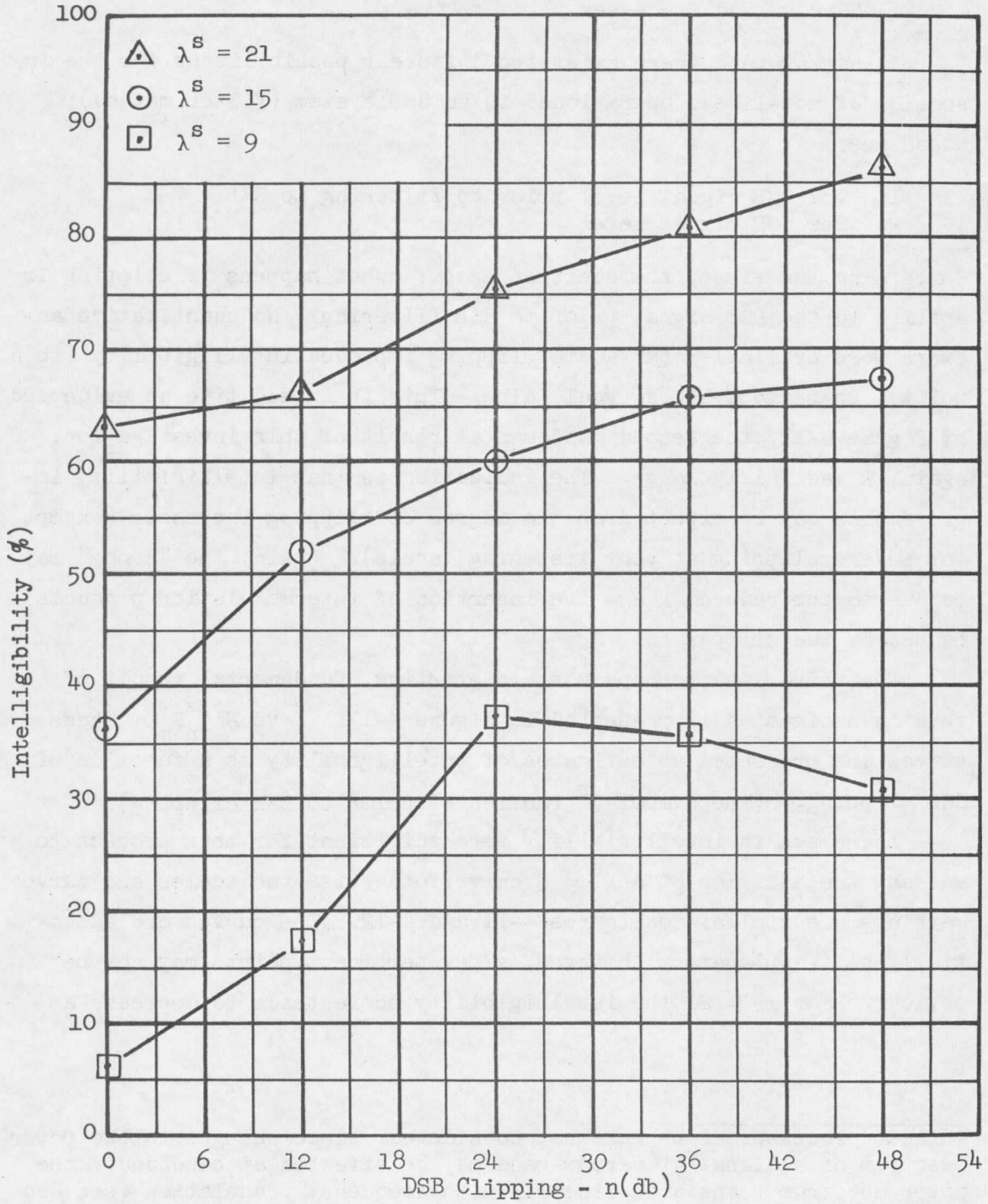


Figure 4-12. Intelligibility Estimates for $\mathbb{P}C_{n\,m}^{d}F^{s}(\lambda^{s})$.

4.2.3. Narrow-band Processes

At narrow-band there exist two different possibilities for the insertion of non-linear operations* in an SSB system (filter method); these are:

1. The DSB signal level prior to filtering to SSB
2. The SSB signal level

Concerning the first, the question arose: what happens if clipping is applied to the DSB signal prior to SSB filtering? No quantitative answers were available; but audio clipping improved intelligibility (to a point), so perhaps $\mathrm{NP}\,C_n^d F_m^s$ would also. This is indeed true as evidenced by Figure 4-12, the second fundamental result of this investigation; again, N was fairly large. The indication is that intelligibility improvements may be expected as the degree of clipping increases (except for severe clipping at poor listening levels). Again, the "droop" may be due to the reduced λ and the insertion of intermodulation products caused by the clipper.

The third, and perhaps most interesting, fundamental result of this investigation is presented in Figure 4-13: I vs $\mathrm{NP}\,C_n^s F_m^s$. These curves are presented as estimates of intelligibility as a function of SSB clipping. (The symbol C^s implies F_m^s prior to SSB clipping).

Increases in intelligibility were sufficient for this process to warrant the addition of a $\lambda^s = 3$ curve; otherwise the scales and curve markings are similar to Figures 4-11 and 4-12. The curves are essentially self-explanatory; however, a few pertinent points may not be obvious. For $\lambda^s = 21$ the intelligibility score tends to decrease as

*Linear processes alone were not considered, since changes in the power spectrum of a signal at narrow-band may be effected at baseband; the power spectrum translates linearly with frequency translation (see Section 2.2.2). This statement is valid for an SSB signal only if $P^a = P^d = 0$.

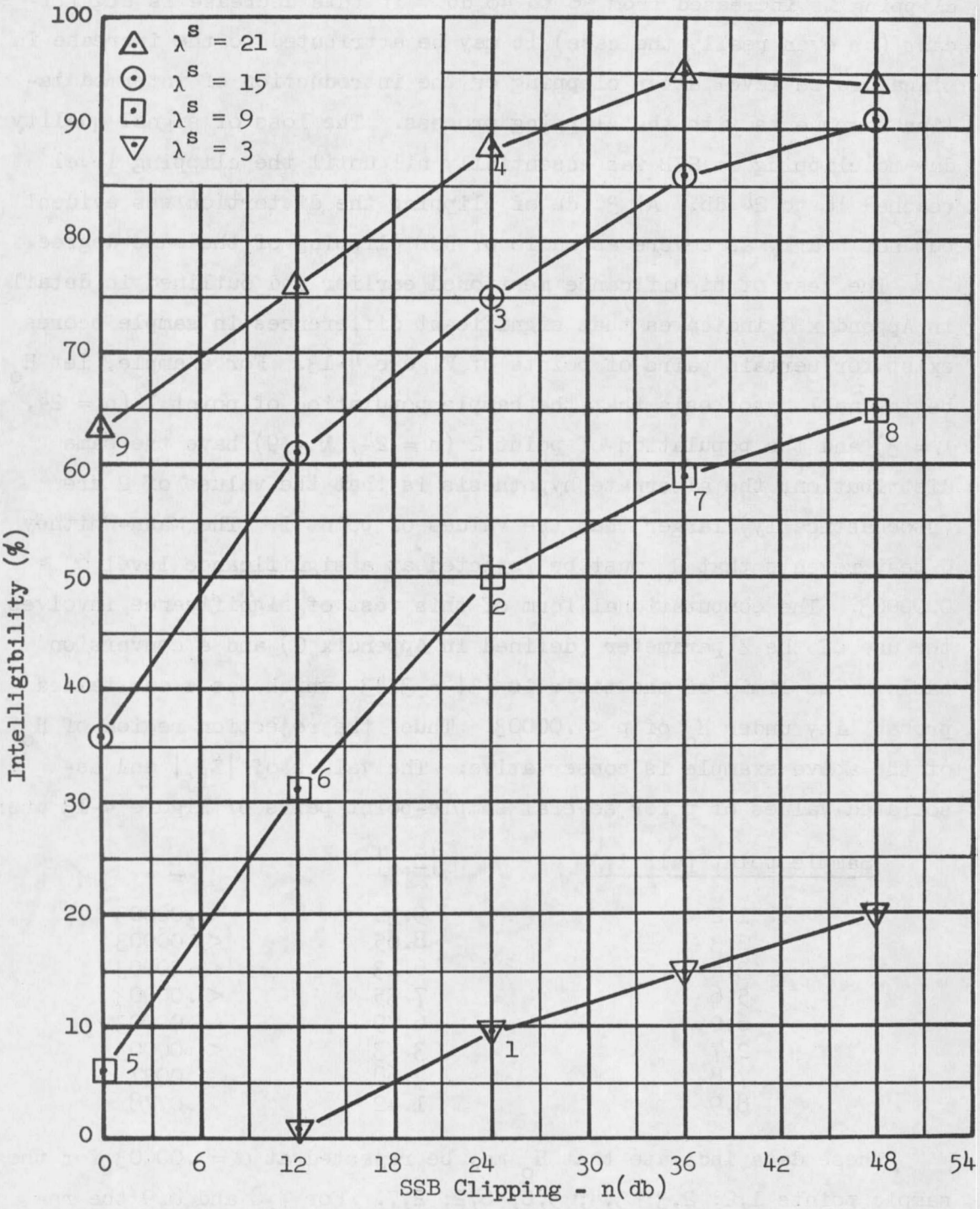


Figure 4-13. Intelligibility Estimates for $\mathbb{P}\, C_n^s F_m^s (\lambda^s)$.

clipping is increased from 36 to 48 db. If this decrease is significant (or even really the case) it may be attributed to the increase in signal noise level after clipping or the introduction of intermodulation components with the clipping process. The loss of signal quality due to clipping at SSB was essentially nil until the clipping level reached 18 to 24 db. At 36 db of clipping the distortion was evident but not nearly as severe as audio or DSB clipping of the same degree.

The test of significance mentioned earlier and outlined in detail in Appendix G indicates that significant differences in sample scores exist for certain pairs of points of Figure 4-13. For example, let H_o be the null hypothesis that the sample population of point 1 ($n = 24$, $\lambda = 3$) and the population of point 2 ($n = 24$, $\lambda = 9$) have the same distribution; the alternate hypothesis is that the values of 2 are (stochastically) larger than the values of point 1. The Mann-Whitney U test reveals that H_o must be rejected at a significance level $\alpha \geq 0.00003$. The computational form of this test of significance involves the use of the Z parameter (defined in Appendix G) and a conversion table. The limit of the table is $|Z| = 3.43$, which has a one-tailed probability under H_o of $p < .00003$. Thus, the rejection region of H_o of the above example is conservative. The values of $|Z_{ij}|$ and associated values of p for several sample-point pairs of Figure 4-13 are:

Sample point pair i,j	$\lvert Z_{ij} \rvert$	p
1,2	9.32	< .00003
2,3	8.65	< .00003
3,4	6.13	< .00003
5,6	7.35	< .00003
6,2	6.79	< .00003
2,7	3.73	< .00003
7,8	2.68	.0037
8,9	1.42	.0778

These data indicate that H_o may be rejected at $\alpha = .00003$ for the sample points 1,2; 2,3; 3,4; 5,6; 6,2; 2,7. For 7,8 and 8,9 the regions of rejection are respectively: $\alpha > .0037$ and $\alpha > .0778$.

The limited tests cited above justify (to an extent) the intuitive evaluation of significant differences derived on the basis of sample (mean) point spacings in the intelligibility domain. The interpretations and discussions which follow, therefore, will make no further reference to tests of significance.

In order to exemplify the use of the data shown in Figure 4-13, let us assume a practical situation:

1. An SSB system has a peak power constraint
2. The channel noise is such that the intelligibility at the receiver output is in the vicinity of 60%.

Now 60% intelligibility may be obtained by using unprocessed speech in the presence of a given amount of noise corresponding to $\lambda = 21$. This same 60% intelligibility may be obtained by using an SSB signal which has been clipped 36 db and has a peak value 12 db below that of the unprocessed signal; that is, 36 db of SSB clipping at $\lambda = 9$ yields the same intelligibility score. Of course, this also assumes white noise as corrupting signal. This situation implies that if the average power constraint of a transmitter is at least 12 db below the peak constraint, the effective transmitter output power can be increased by 12 db (approximately) if the SSB signal is clipped 36 db prior to transmission. Similar conclusions may be drawn by considering other segments of Figure 4-13. For an unprocessed signal at $\lambda = 15$, the intelligibility score is approximately 36%. Maintaining the same peak value and the same amount of additive noise, the addition of 36 db of SSB clipping yields an intelligibility score of approximately 86%.

Thus far, audio, DSB, and SSB clipping as applied to an SSB system have produced, at least in some ranges, increased intelligibility scores for constant λ. It might be surmised at this point that combinations of these processes might have an advantage over any single process. Further, these data have been given for signals which had $\lambda \approx$ 45 db prior to clipping; hence an interesting consideration might be: How far do the intelligibility estimates of Figures 4-11, 4-12, and 4-13 hold for a more-noisy, clipped signal? The implications of these

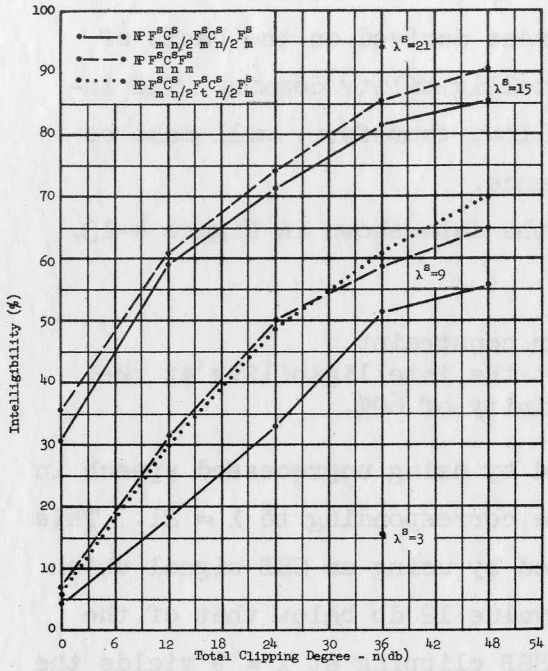


Figure 4-14. Intelligibility Estimates for Iterative SSB Clipping: A Comparison.

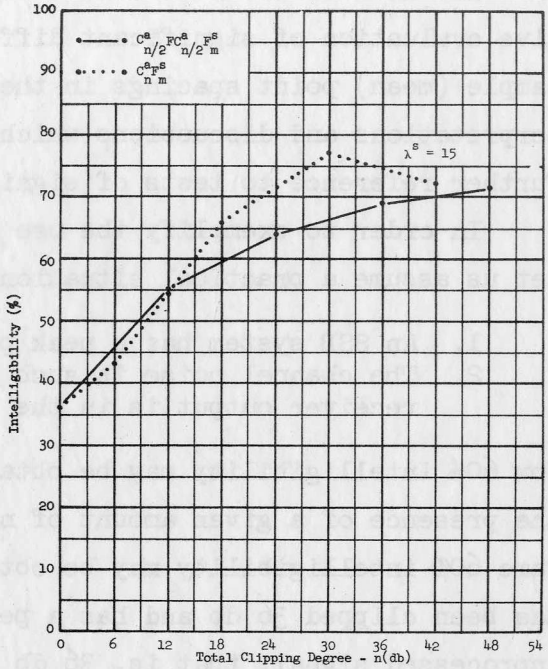


Figure 4-15. A Comparison of Single and Iterative Audio Clipping (λ^s = 15).

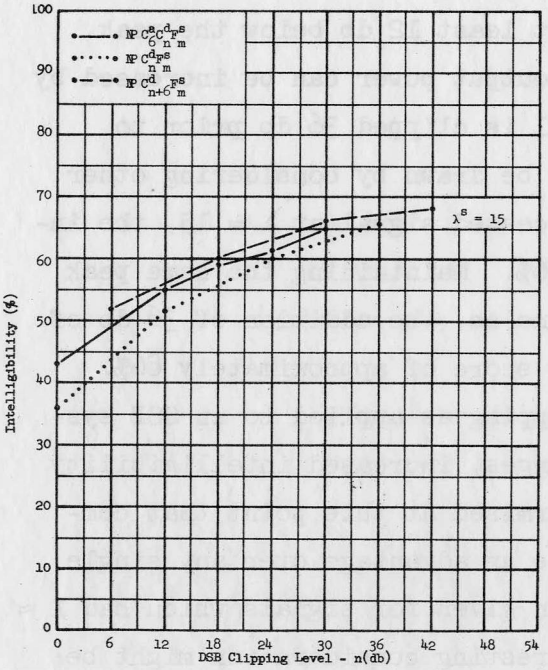


Figure 4-16. Interdependencies Between Audio and DSB Clipping (λ^s = 15).

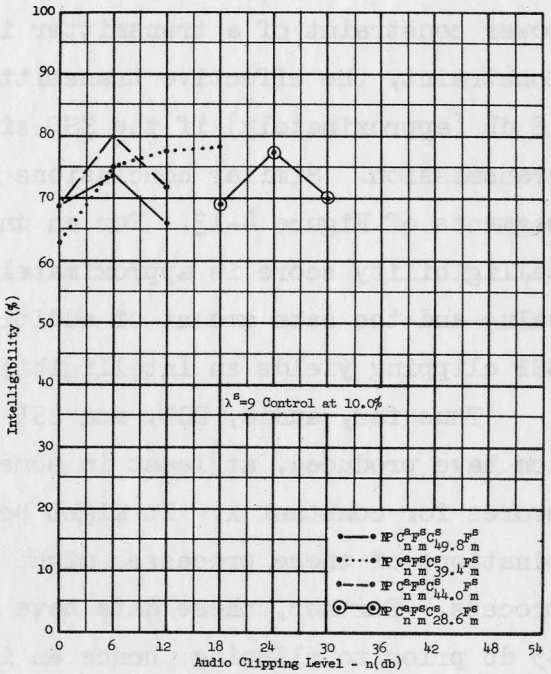


Figure 4-17. Interdependencies Between Audio and SSB Clipping (λ^s = 9).

statements and the effects of spectrum tilts prior to clipping will be considered in later sections.

4.2.4. Iterative, Interdependent and Other Combinatorial Processes

4.2.4.1. Iterative Processes

In practice it is difficult to obtain over 24 db of narrow-band clipping in one operation. This physical restriction gives rise to an increase in latitude of clipping operations at SSB. For example, the solid curves of Figure 4-14 provide intelligibility estimates for $\lambda^s = 9, 15$, when two narrow-band clippers are isolated with a mechanical filter. The first clipping operation produces a flat-topped wave, peaking occurs in the filter, the resulting wave is again clipped; a reduction to bandwidth in the last mechanical filter precedes signal introduction into the channel. The dashed curves of Figure 4-14 are the $\lambda = 9, 15$ curves of Figure 4-13; these are included for comparison purposes. Note that the total amount of clipping for corresponding data points of these curves is the same. For the $\lambda = 15$ case, there is a small difference in intelligibility estimates; however, in the $\lambda = 9$ case, the difference of intelligibility scores between the two processes is more marked. The points marked $\lambda = 3$ and $\lambda = 21$ are associated with the iterative-clip process with a mechanical filter interposed. The dotted curve of this figure provides an indication of intelligibility scores when the mechanical filter between clippers of the former case is replaced by a buffered tuned-circuit, with a Q of approximately 37 at 250 kc. This curve closely resembles the curve for conventional clipping.

The iterative-clip process for the audio case at $\lambda^s = 15$ is given in Figure 4-15 by the solid curve. The dotted curve of this figure is the $\lambda^s = 15$ curve from Figure 4-11. These data imply, as in the preceding case, that no advantage is gained by "splitting up" the clipping process.

A series of indicator tests using small N ($N \approx 10$) was run using unequal values of clipping in the iterative-clip process for SSB. For example, with the first clipper adjusted for n = 6, the second clipper was adjusted for n = 18, 24, 30, 36, (all $\lambda^s = 9$); no improvements in intelligibility scores were noted over the case where the clipping was performed in one step. Similar tests for the first clipping level at 12 and 18 db indicated similar results.

4.2.4.2. Interdependent Processes

Earlier it was shown that audio, DSB and SSB clipping, when considered individually, increased intelligibility scores in major portions of the clipping spectra presented. Since each of these clipping processes yielded a favorable result, this section will be devoted to the consideration of the interdependencies that may exist between the clipping processes at these levels. Figure 4-16 depicts such a result; the solid curve shows intelligibility estimates for various degrees of DSB clipping preceded by 6 db of audio clipping. The dotted curve of Figure 4-16 is the $\lambda^s = 15$ curve from Figure 4-12; that is, the intelligibility estimate for $\lambda^s = 15$ and n db of DSB clipping only. The dashed curve is also taken from Figure 4-12, except that in this case the total clipping level is the same as for the case of 6 db of audio pre-clipping. The implication of this figure is that 6 db of audio pre-clipping increased intelligibility scores over the case where preprocessing was not used; however, the estimates did not appear to equal those obtained when the total clipping operation was performed at DSB.

A series of intelligibility indicator tests were run to determine the interdependencies that existed between DSB and SSB clipping in $\mathrm{NP}\, C^{d}_{n} F^{s}_{m} C^{s}_{n} F^{s}_{m}$ for $\lambda^s = 9$. Tests were run for n = 0, 6, 12, 18, 24 db. However, in all cases the intelligibility estimates ranged significantly below the intelligibility scores obtained if the total clipping level was achieved at SSB. Neither graphical nor tabular data are presented for these tests.

Figures 4-17, 18, 19 point out the interdependencies that exist between audio and SSB clipping. Figure 4-17, for example, shows parametric curves of constant level of SSB clipping with audio clipping-level as the independent variable. Note that, in three of the four parametric curves shown, an apparent maximum is produced: for $C^s_{28.6}$ (solid curve, circle identification) the maximum occurs at an audio clipping level of 24 db, making a total clipping level of approximately 53 db. The apparent maximum for $C^s_{49.0}$ appears at 6 db of audio clipping, making a total clipping of approximately 56 db. The maximum for $C^s_{44.0}$ occurs at in the vicinity of 6 db of audio clipping for a total of approximately 50 db clipping. It appears that the curve for $C^s_{39.4}$ is approaching a maximum at perhaps 18 db, making a total clipping level of approximately 57 db. Note in these cases that the intelligibility scores begin to fall as the total clipping level is increased above certain levels, with these levels bracketed in the range of 50 to 57 db (approximately). One possible reason for this behavior is that such clipping levels actually represent clipping into the noise level of the system signals (recall that the λ of the test tapes was of the order of 45 db).

The reason for the rather odd clipping levels in the above analysis is that, after the test tapes had been recorded, an error in clipper calibration was discovered. The clipping error constants were determined for each clipping level, and the corrections were made on the parametric curves rather than to re-tape test materials (the recording of the test tapes for these interdependency evaluations was particularly time-consuming).

The information contained in Figure 4-18 is similar to that of Figure 4-17, except that $\lambda^s = 15$. In this figure the parametric curves are for SSB clipping of 16.6, 28.6, and 40.6 db. Note that a maximum was reached for the 40.6 curve at an audio clipping level of approximately 12 db; a total clipping level of approximately 53 db (again). The other two curves did not yet reach an indicated maximum

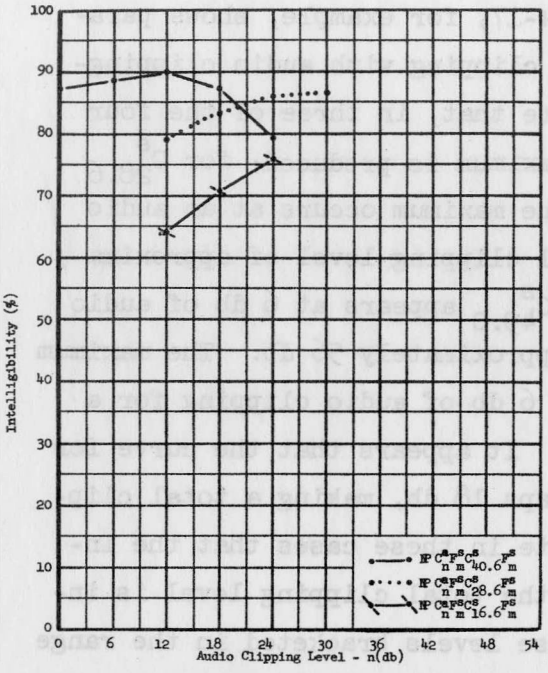


Figure 4-18. Interdependencies Between Audio and SSB Clipping ($\lambda^s = 15$).

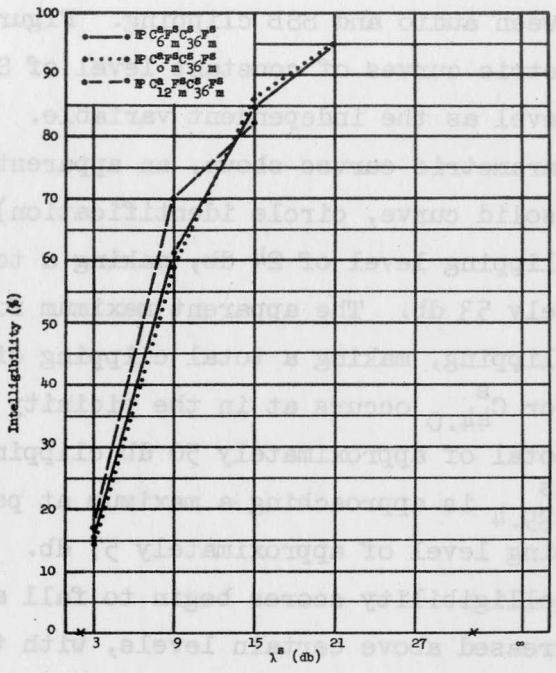


Figure 4-19. The Effect of Audio Pre-Clipping on $C^s_{36}F^s_m$.

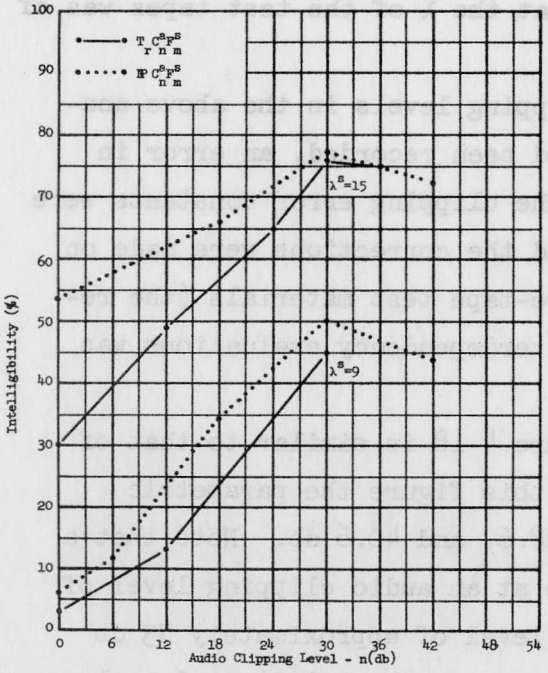


Figure 4-20. The Effect of T_r Pre-Processing on $C^s_{nm}F^s$ (λ^s = 9,15).

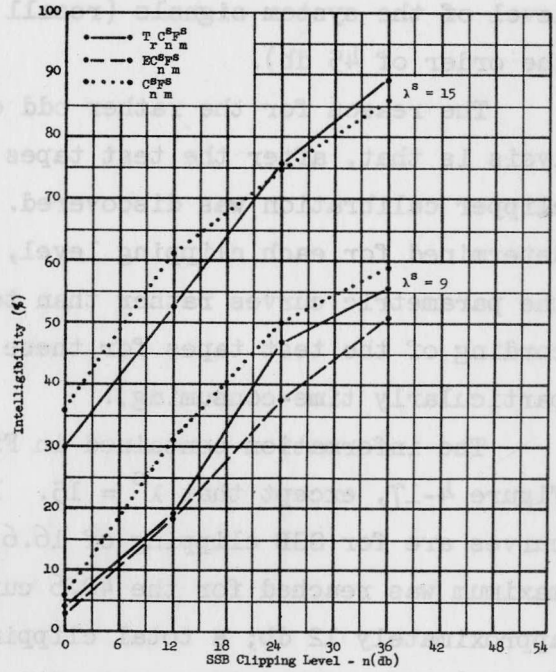


Figure 4-21. The Effect of T_r and E Pre-Processing on C^s_n (λ^s).

and additional data points could have been taken in order to complete the analysis of the maxima-trend.

The effects of 0, 6, and 12 db of audio pre-clipping to 36 of SSB clipping are shown in Figure 4-19. The dotted curve (with 0 audio clipping) was taken from Figure 4-13 for n = 36. Figure 4-19 indicates that, for relatively severe SSB clipping, an audio pre-clipping of 6 db reflects essentially no change in the intelligibility scores, while 12 db of audio pre-clipping increases the intelligibility estimates at the lower λ levels.

The data of this section seem to point out that most iterative processes are not as advantageous as a single clipping operation (particularly when that single operation is performed at SSB).

4.2.4.3. Linear Pre-Processing

In the early studies of this investigation it was found that differentiation of the speech signal prior to premodulation clipping had no advantageous effect on intelligibility scores when λ was maintained at audio (with 5 kc bandwidth). The same result is depicted in Figure 4-20, when $\lambda^s = 9$, 15 is considered. The indication is that for relatively low degrees of clipping, the addition of the differentiation process may offset the effects of as much as 15 db of audio clipping (consider $\lambda^s = 15$ at an intelligibility score of 54%).

Figure 4-21 shows the effects of differentiation and equalization as pre-processes on the SSB-clip series. The solid curves of this figure represent differentiated signals. The dashed curve represents an E pre-process for $\lambda^s = 9$, and the dotted curves are for the indicated amounts of C^s with no pre-processing. For $\lambda^s = 9$ the intelligibility estimates for both T_r and E fall below the estimates of C^s alone. For $\lambda^s = 15$ and the lower SSB clipping levels, the T_r process did not yield an intelligibility gain; however, at the higher clipping levels (and consequently better listening conditions) the T_r pre-process indicates a gain over straight clipping. This is the only point where this

phenomenon occurred during the course of this investigation.

The effect of the T_r pre-process on iterative clipping is shown in Figure 4-22 for various λ levels. For this case, T_r appears to offer no advantage.* The data for the dotted curve of Figure 4-22 were taken from Figure 4-19.

4.2.4.4. Reduced Signal λ

All of the preceding tests were conducted using test materials with an unprocessed λ of approximately 45 db. Since all signals in physical systems do not have $\lambda = 45$ or better, it was decided to run a brief indicator series of tests to determine the effects of a reduced signal λ condition.

In order to investigate these effects, white noise was added to the materials from the original test tapes to an extent that the resulting signal had $\lambda = 27$ db. The resulting purposely-corrupted signal was recorded as a reference; these materials hereafter will be referred to (as though they had been processed) as λ^{27}. These "new" test materials were then used as the system input signal; the degrees of clipping and channel λ were defined using the peak values of this "new" signal.

For the case of audio clipping (λ^s), the λ^{27} condition showed an appreciable decrease in intelligibility estimates for given degrees of clipping: this is indicated in Figure 4-23. The solid curves are the

*The process of differentiation emphasizes signal high-frequency components; this tends to make the speech sound more "crisp". When differentiation precedes severe clipping, this crispness does not appear desirable since the clipping process followed by bandpass filtering also tends to make the speech sound more "crisp" and "tinny". It may be that this increase in the "tinny" quality of the speech signal resulting after differentiation and clipping (regardless of where the clipping was performed) was a contributory factor in the decrease of intelligibility scores.

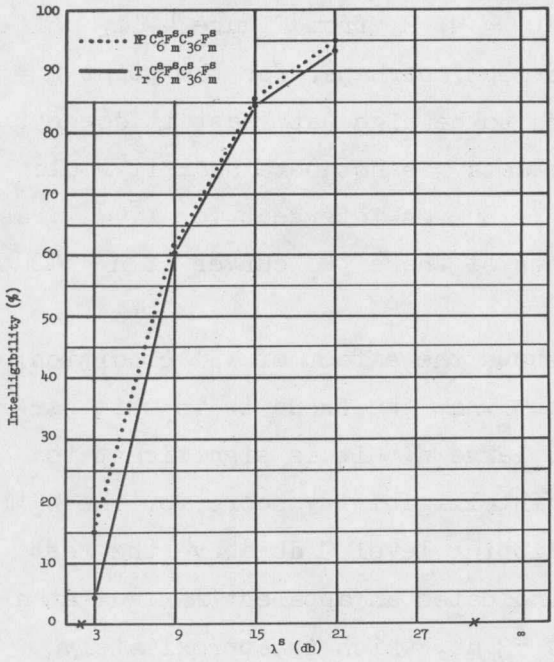


Figure 4-22. The Effect of T_r on $C^a_6 F^s_m C^s_{36} F^s_m$

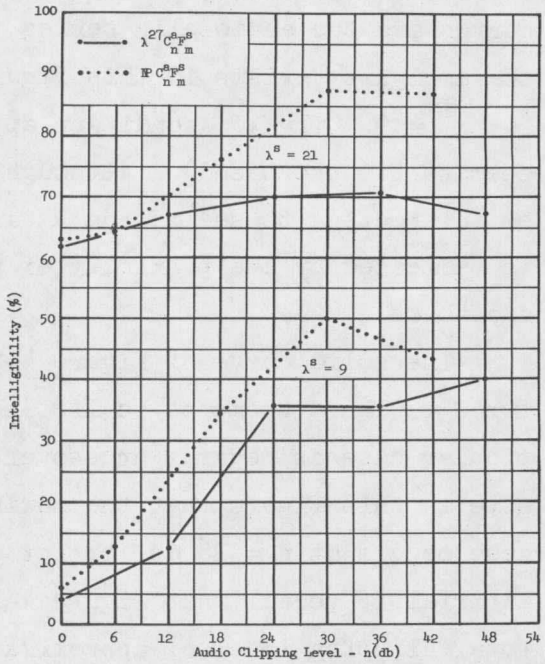


Figure 4-23. The Effect Upon I of Reducing Signal λ Prior to Audio Clipping (λ^s).

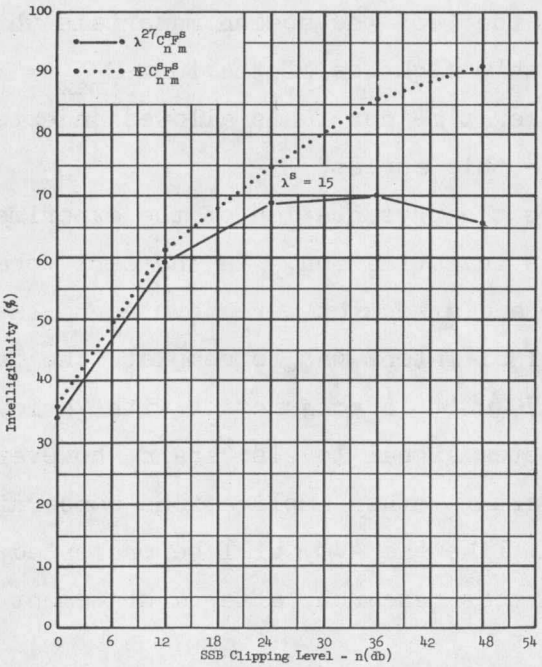


Figure 4-24. The Effect Upon I of Reducing Signal λ Prior to SSB Clipping ($\lambda^s = 15$).

λ^{27} conditions followed by n db of audio clipping, while the dotted curves are the audio-clip series for λ^{s} = 9, 21 from Figure 4-11. Little ambiguity exists in this figure except, perhaps, for the point n = 24, λ^{s} = 9. The λ^{27} condition appears to be high (at least it doesn't seem to fit the trend). Although the data was not obtained, it would be interesting to see if the NP and λ^{27} curves intersect for λ^{s} = 9, as is indicated by the proximity of points of these two curves at n = 42 and 48 db.

The solid curve of Figure 4-24 shows the effect of λ^{27} condition for C_n^s. There seems to be little doubt what the trend is in this particular case as regards crossovers at large n. It is significant to note in this figure that the maximum intelligibility score for the λ^{27} case occurs at n = 36 db, or for a clipping level 9 db above the test material λ; recall that Figure 4-18 indicates an apparent maximum at a total clipping level of approximately 53 db, which is approximately eight db above the λ of the test materials. This might indicate that there exists a parametric relationship between the maximum permissible clipping level and the λ of the speech materials which are being clipped; this relationship might be of the form $C_{max} = \lambda + \delta$, where $\delta \approx 8$ db. However, neither time nor funds allowed an extension of the investigation to include this aspect.

This completes the presentation of the experimental data obtained as a result of this investigation. In the early phases, a considerable amount of time and attention were devoted to the aspects of an AVC unit, and the original intent was to complete the data series with an inclusion of the AVC unit in conjunction with various degrees of clipping at various system signal levels; again, however, time and funds did not permit. Certain other implications concerning the SSB processing parameters, like the AVC, will be presented in the section on suggestions for further research, after a statement of conclusions.

5. DISCUSSION; INTERPRETATIVE LIMITATIONS

The interpretation of data as set forth in the previous chapter can best be classified as reasonable estimates of the intelligibility scores associated with several signal processing techniques. The estimates are based on a summation of data gathered from a number of trials. In some cases, the intelligibility estimates were based on relatively large number of trials; the outstanding examples in this category are data presented for controls, audio clipping, DSB clipping, and SSB clipping (all with λ determined at narrow-band). Other intelligibility tests were provided as indicators of trends or ranges, and the probability of error for these later tests is significantly higher than the tests of large N, as evidenced by certain data points which appeared to have extraneously high or low values.

The major sources of error encountered in the acquisition of the experimental data were:

1. Equipment
2. Recording
3. Subject variation
4. Scoring

Obviously, the contributory effects of these errors were minimized as much as possible; the preventative measures which were taken will be reviewed here.

The laboratory models of the SSB system and processing devices provided little cause for concern except for clipper calibration. As previously noted, clipper calibration was in error for one set of data; this error was readily discovered, however, by means of calibration checks which were applied periodically. The frequency standard and noise generator were allowed to operate continuously so that warm-up and stabilization procedures were not necessary. The remainder of the equipment was allowed a warm-up period of approximately 30 minutes prior to recording of test tapes. Minor difficulties were encountered with the source tape recorder in that the output level would change in

a random fashion; this was corrected through routine maintenance. Periodic checks of speed, equalization, flutter and wow were performed on each tape recorder to ensure that these characteristics were within manufacturer's specified limits.

The errors due to test setup and recording were minimized by having two project members present during recording sessions. One member made the necessary inter-connections of system components for a particular processing operation, and the other member checked the setup prior to recording. The quantity ΔPV provided an invaluable tool for insight into possible recording errors.

Errors due to listener variations were minimized by:

1. Subject selection based on hearing ability, and the temporary rejection of a subject with a cold.
2. The training period.
3. Minimization of distraction
   a. Variable volume level was used.
   b. The noise was presented at a fixed level and changes in λ were made by changing the signal level.
   c. Only female team members were used.
4. Minimization of boredom by break periods.
5. Process, λ, and list randomizations.

Minimization of errors due to scoring and data reduction was accomplished by obvious methods: the test was scored immediately following the test run, no time limit was placed on the scoring team, and the test supervisor spot-checked all scoring.

Any interpretation or extrapolation of the data presented in this report must be done with care, since the framework of the study is restrictive. A summary of these restrictions are:

1. Bandwidth was 3 (or 5) kc
2. Subjects were well-trained listeners
3. Single-word stimulus materials were used to obtain intelligibility estimates
4. λ definition of signal-to-noise was used
5. A peak-limited SSB system was assumed

Even considering these restrictions, perhaps additional ramifications and implications are extractable from the findings presented in this report; however, such extensions are not obvious at this time.

6. RESULTS AND CONJECTURE

In the explicative phases of this research, certain procedures, designs and definitions were presented which may be of value in themselves. Included in this category is the information regarding:

1. The articulation test
   a. Test materials
   b. Tape preparation
   c. Team selection
   d. Scoring aspects
2. The definition of λ, which is suitable for peak-limited systems
3. The method of clipper representation given in Appendix D-2
4. The general equipment and experimental design philosophy

A review of the signal processes as applied strictly to audio signals was presented. The first result was the equivalence of intelligibility data as obtained by the fixed and variable volume levels.

The linear processes E and T_r provided minor increase in intelligibility estimates for λ^a; for λ^s, and in conjunction with nonlinear operations at audio, DSB and SSB; these linear operations generally gave rise to a signal intelligibility degradation.

The audio AVC unit provided nominal increases in intelligibility, and it is unfortunate that a combination of the AVC-clip process could not have been investigated. The transition from a 5 kc bandwidth to a 3 kc bandwidth provided additional insight into the definition of λ.

The three most fundamental (and most accurate) intelligibility estimates were made for: audio, DSB, and SSB clipping with λ defined at narrow-band. Of these three SSB clipping provided the most significant intelligibility increases and at the same time, more fully preserved signal quality. The interdependent and iterative combinations of these processes yielded advancements in the intelligibility scores; however, in no case did the intelligibility estimate exceed that derived from the total amount of processing provided at SSB.

The presentation of results was completed by the description of a reduced-λ signal condition and its effect on the intelligibility estimates of audio and SSB clipping. It was noted that there appeared to be a constant relationship existing between maximum allowable clipping and the λ of the unprocessed test materials: $C_{max} = \lambda + \delta$, $\delta \approx 8$ db.

In the process of gathering the foregoing data, the author has gained considerable insight into the design and construction of actual equipments and tests used in the determination of the intelligibility estimates. Such new-found experience and knowledge gives rise to a certain dissatisfaction with the results (both positive and negative) obtained, and tend to lead to implications of ideas that could have or should have been considered. Some of these ideas may be of value, some may be ridiculous and some may be trivial; however, in general, the ideas fall into the category of suggestions for further research.

Those directly related to this study are:

1. The completion of the interdependency series may provide valuable insight into the clipping operation in general.
2. The consideration of both open- and closed-loop AVC units with both audio and SSB control may be of value. For example, suppose an AVC unit had control upon the amplitude of the audio input signal, but was controlled by the envelope function of the narrow-band SSB signal.
3. In close relationship to the AVC operation, perhaps a tone of fixed frequency could be added to a signal prior to modulation with amplitude inversely proportional to the amplitude of the narrow-band modulating function with a view toward reducing peak-factor.
4. The linear operations of concern in this investigation afforded little, if any, increase in intelligibility estimates; however, other linear filtering techniques may provide an advantage.
5. Are the processing techniques outlined in this investigation still useable when frequency translational error is present in addition to a corruptive noise signal?
6. The quadrature function method of SSB generation provides interesting possibilities for non-linear signal operations which are not present in the more standard SSB systems.

The ideas of a more general or a more speculative nature might include:

1. Phasing or autocorrelation techniques may be a possibility for the reduction of the peak-factor of speech.
2. Define the lower limit of bandwidth with suitable processing for "acceptable" speech.
3. Investigate SSB processing techniques as applied to satellite communication systems for both aural and digital data.
4. An investigation of signal processing techniques applicable to other types of transmission (i.e., DSB, FM, FSK, etc.) might be considered.
5. An investigation of speech processes of an unconventional nature which may arise as a result of using a vocoder.
6. Investigate a signal peak detection, delay and addition procedure so that the resulting signal has a peak value less than normal speech (for a given average value).

Some of these ideas are indeed conjecture and must remain so until the structures of the speech processing complex are more fully defined. It is to be hoped that this paper provides a starting point for a more specific study in one of these areas.

7. APPENDICES

Appendix A - Literature Consulted

1. I. B. Crandall and D. MacKenzie, "Analysis of the Energy Distribution in Speech", Physics Review, 19, 221, (1922).

2. E. E. David, M. V. Mathews, H. S. MacDonald, "Description and Results of Experiments with Speech Using Digital Computer Simulation", IRE Wescon Convention Record, Part 7 (1958).

3. E. E. David, "Artificial Auditory Recognition in Telephony", IBM Journal Res. & Dev., 2, (1958).

4. H. W. Dudley, "The Vocoder", Bell Labs Record, 18, 122, (1939).

5. J. P. Egan and F. M. Wiener, "On the Intelligibility of Bands of Speech in Noise", JASA, 18, 435, (1946).

6. J. P. Egan, etal, "Studies on the Effect of Noise on Speech Communication", Psycho-Acoustic Laboratory, OSRD Report #2038, (RB22907), (1943).

7. N. R. French and J. C. Steinberg, "Factors Governing the Intelligibility of Speech Sounds", JASA, 19, 90, (1947).

8. L. R. Kahn, "The Use of Speech Clipping in Single-Sideband Communications Systems", Proc. IRE, 45, 1148, (1957).

9. V. O. Knudsen, "The Hearing of Speech in Auditoriums", JASA 1, 56, (1929).

10. J. C. R. Licklider, "Effects of Amplitude Distortion Upon the Intelligibility of Speech", JASA, 18, 429, (1946).

11. J. C. R. Licklider and I. Pollack, "Effects of Differentiation, Integration and Infinite Peak Clipping Upon the Intelligibility of Speech", JASA, 20, 42, (1948).

12. D. W. Martin, "Uniform Speech Peak Clipping in a Uniform Signal-to Noise Spectrum Ratio", JASA, 22, 614 (1951).

13. W. H. Martin, "Transmitted Frequency Range for Telephone Message Circuits", BSTJ, 9, 483, (1930).

14. L. E. McCoy, "Articulation Testing Facility at Montana State College", ERL Technical Report, Montana State College, Bozeman, Montana, (1958).

15. J. F. Nickerson, "The Effect of Frequency Translational Error on the Intelligibility of Speech in the Presence of Noise", ERL Technical Report, Montana State College, (1958).

16. Nickerson, Miller, Shyne, "A Comparison of Five Articulation Tests", ERL Technical Report, Montana State College, (1960).

17. I. Pollack, "Effects of High-Pass and Low-Pass Filtering on the Intelligibility of Speech in Noise", JASA, 20, 259, (1948).

18. N. A. Shyne, "An Investigation of Selected Signal Processing Techniques", ERL Technical Report, Montana State College, Bozeman, Montana, (1960).

19. S. Siegel, "Nonparametric Statistics for the Behavioral Sciences", McGraw-Hill Book Company, New York (1956).

20. W. W. Smith, "Premodulation Speech Clipping and Filtering", QST, 30, 46 (1946).

21. J. C. Steinberg, "Effects of Distortion on the Recognition of Speech Sounds", JASA, 1, 121, (1929).

22. J. Q. Stewart, "An Electrical Analog of the Vocal Organs", Nature, 110, (1922).

23. W. Wathen-Dunn and D. W. Lipke, "On the Power Gained by Clipping Speech in the Audio Band", JASA, 30, 36, (1958).

24. D. K. Weaver, "Quadrature Signal Functions and Applications", ERL Technical Report, Montana State College, (1958).

Appendix B - Signal Types

Signals in physical systems are band-limited. Certain classes of signals are defined with respect to significant spectral components. A low-pass signal is one containing frequency components between zero frequency and some upper limit f_2. A band-pass signal is defined as one whose significant frequencies lie between two frequencies f_1 and f_2, where f_1 may be arbitrarily small but not equal to zero (a band-pass signal has no dc component).

The band-center and bandwidth of a signal are, respectively;

$$f_o = \frac{f_1 + f_2}{2} \quad , \quad W = f_2 - f_1$$

A broad-band signal is a band-pass signal which satisfies the condition that $W \approx f_2$, while, for a narrow-band signal, W is much less than f_o. Baseband speech is classified as a broad-band signal.

The above signal definitions are described pictorially in Figure B-1.

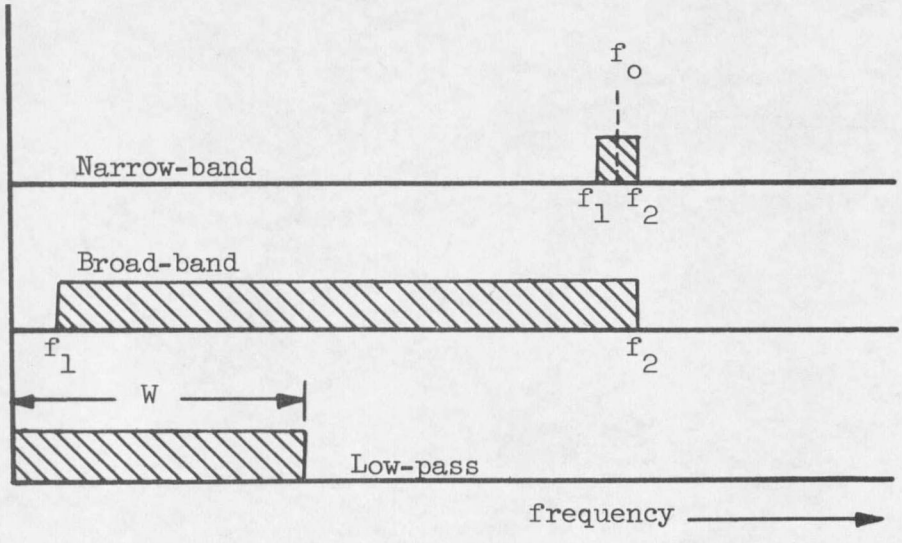


Figure B-1. Band-Limited Signals

Appendix C - Model SSB Systems

In order to implement laboratory investigation of system and signal characteristics in conjunction with processing operations, model SSB generators of the filter and quadrature function types were constructed.

1. Filter Method

Figure C-1 shows the components used to derive the SSB signal using the filter method.* The input and output impedances of the units were 600 ohms (where applicable) to facilitate interconnection and insertion of processing devices. The DSB generator was followed by a mechanical filter amplifier to produce SSB. Collins 250Z5 (250 kc, lower sideband, 2.7 kc bandwidth) filters were used. A typical response and phase characteristic for these filters is given in Figure 2-6, and a pictorial study of a system test is presented in Figure 3-2.

The receiver and frequency standard (a stable crystal oscillator) were usable at 100 and 250 kc. The output of the DSB generator had a carrier suppression of approximately 40 db. The DSB generator and SSB receiver translation frequencies were derived from the same standard to eliminate frequency translational error. Purposeful translational error could be introduced by means of an interpolation oscillator used in conjunction with the frequency standard.

2. QF Method

For an explication of the QF method of SSB generation, reference is made to Figure C-2. If a series representation for the input is used,

---

*For the circuit diagrams of this report, all resistors are 1/2 watt and all capacitance values are expressed in μf unless otherwise specified.

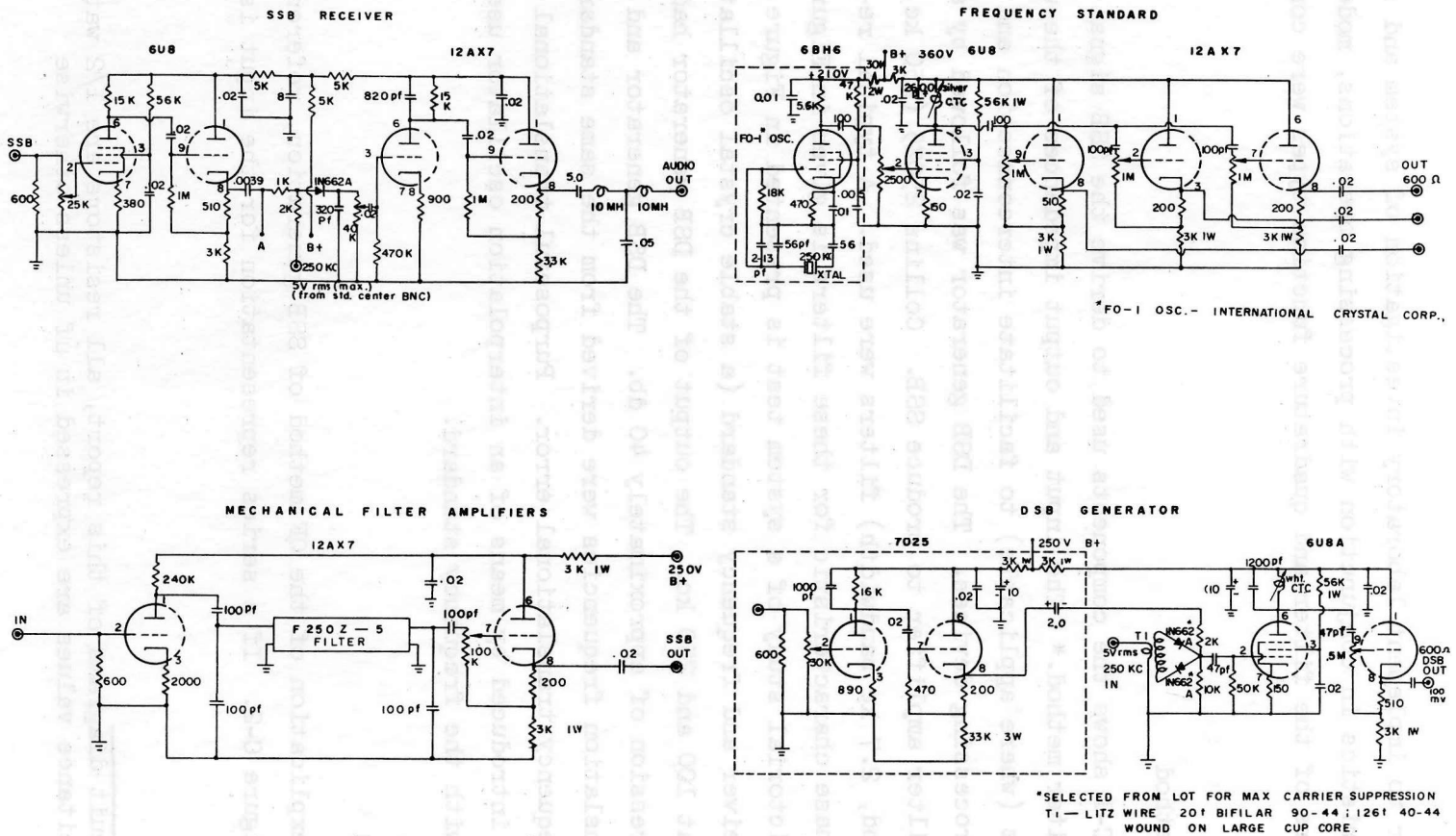


FIGURE C-I. LABORATORY SSB SYSTEM COMPONENTS

$$e(t) = \Sigma\ a_i\ \cos(\omega_i t + \phi_i)$$

the respective outputs of balanced modulators 1 and 2 are:

$$K_1\ [\sin\ \omega_1 t]\ [\Sigma\ a_i\ \cos\ (\omega_i t + \phi_i)] =$$

$$\tfrac{1}{2}K_1\ \Sigma\ a_i [\sin\ (\omega_1 t + \omega_i t + \phi_i) + \sin\ (\omega_1 t - \omega_i t - \phi_i)]$$

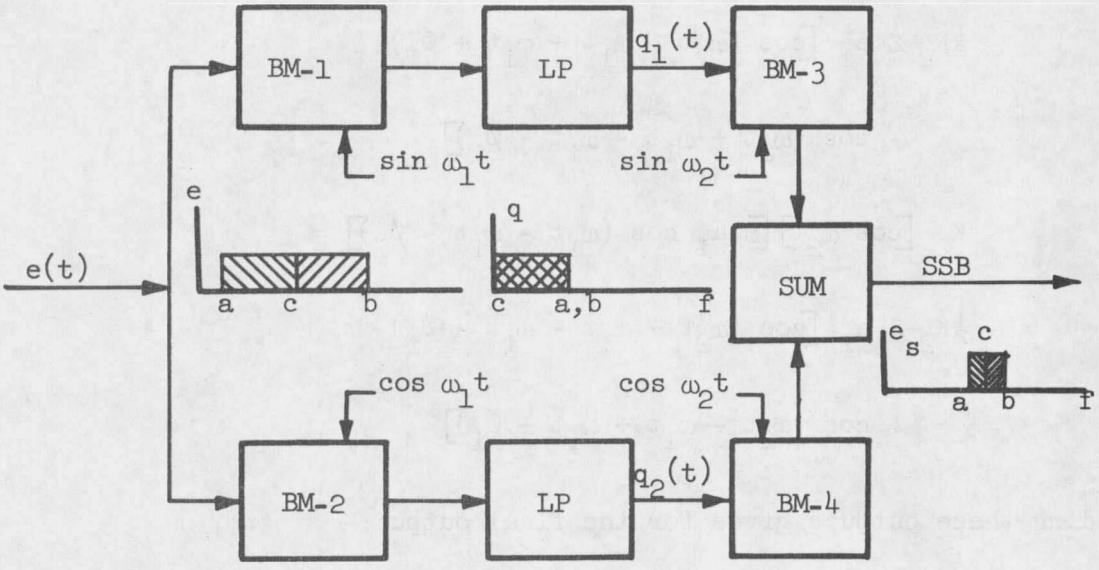


Figure C-2. Quadrature Function Method of SSB Generation

$$K_1\ [\cos\ \omega_1 t]\ [\Sigma\ a_i\ \cos\ (\omega_i t + \phi_i)] =$$

$$\tfrac{1}{2}K_1\ \Sigma\ a_i\ [\cos\ (\omega_1 t + \omega_i t + \phi_i) + \cos\ (\omega_1 t - \omega_i t - \phi_i)]$$

where ω_1 is the baseband center frequency.

The low-pass filters eliminate the sum frequencies in each channel, so that the inputs to balanced modulators 3 and 4 are respectively:

$$q_1(t) = K_2 \Sigma a_i \sin (\omega_1 t - \omega_i t - \phi_i)$$

$$q_2(t) = K_2 \Sigma a_i \cos (\omega_1 t - \omega_i t - \phi_i),$$

which are the quadrature functions. When these inputs are mixed with the second converting frequency (ω_2) with phase as shown, the balanced modulator outputs are:

$$K_2 \left[\sin \omega_2 t\right] \left[\Sigma a_i \sin (\omega_1 t - \omega_i t - \phi_i)\right] =$$

$$\tfrac{1}{2} K_2 \Sigma a_i \left[\cos (\omega_2 t - \omega_1 t + \omega_i t + \phi_i)\right.$$

$$\left. - \cos (\omega_2 t + \omega_1 t - \omega_i t - \phi_i)\right]$$

$$K_2 \left[\cos \omega_2 t\right] \left[\Sigma a_i \cos (\omega_1 t - \omega_i t - \phi_i)\right] =$$

$$\tfrac{1}{2} K_2 \Sigma a_i \left[\cos (\omega_2 t + \omega_1 t - \omega_i t - \phi_i)\right.$$

$$\left. + \cos (\omega_2 t - \omega_1 t + \omega_i t + \phi_i)\right]$$

Adding these outputs gives for the final output:

$$s(t) = K_2 \Sigma a_i \cos \left[(\omega_2 - \omega_1 + \omega_i)t + \phi_i\right] ,$$

which is the original band of frequencies shifted up by an amount ($\omega_2 - \omega_1$). This original band, limited by the low-pass filters, is now centered about ω_2.

A circuit diagram for the QF SSB generator is shown in Figure C-3. The audio input signal is applied to the first set of balanced modulators by a bifilar-wound transformer. Conventional ring-type diode modulators were used for this set. The first translation frequency (2500 cps) is obtained from a divider chain. The 45° phase-shift

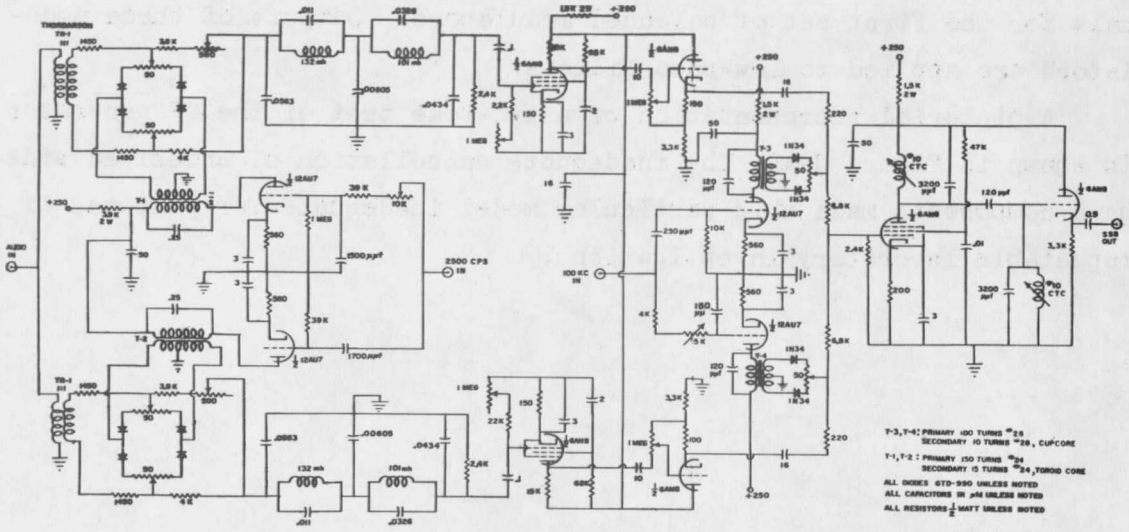


FIGURE C-3, QF SSB GENERATOR

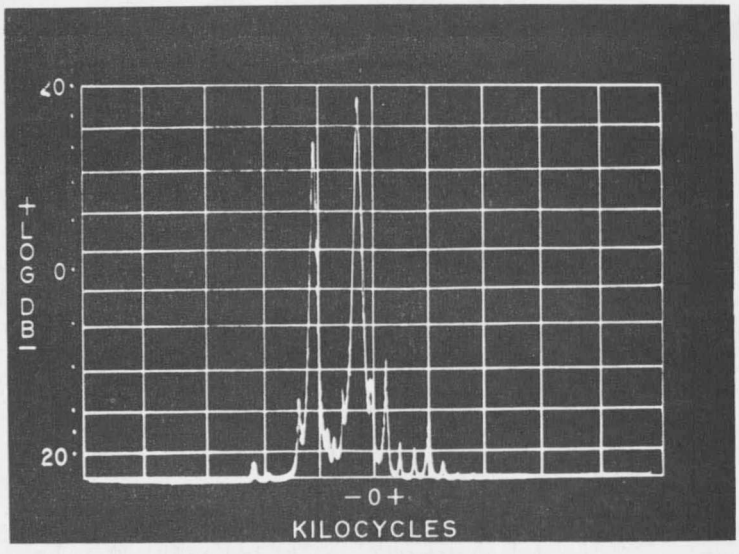


Figure C-4. Spectral Study of QF Generator

networks (of opposite shift) are used to obtain signals in quadrature at the plates of the 12AU7; these signals are used as switching signals for the first set of balanced modulators. Outputs of these modulators are applied to low-pass filters.

A pictorial representation of a two-tone test of the QF generator is shown in Figure C-4. The inadequate cancellation of undesired sideband components made this particular model inadequate for purposes of repeatable laboratory investigations.

Appendix D - Processing Devices

For clarity and completeness, total circuit diagrams and additional pertinent information regarding the speech processing devices are presented here.

1. Clipper

A basic passive clipper as shown in Figure D-1a was used for peak clipping of speech, DSB, and SSB signals.

When all diodes conduct, the output is shorted to the input (ideally), and $e_i = e_o$ as shown in Figure D-1b.

As e_i increases positively, current through the matched resistors, R_1, and diodes D_2 and D_3 increases until D_1 and D_4 become reverse biased and cut off. When D_1 and D_4 cut off, the equivalent circuit reduces to that of Figure D-1c. The output is now equal to a constant value

$$e_o = \frac{VR_2}{R_1 + R_2},$$

and the input and output are isolated. Since $e_o = e_i$ until saturation occurs and the boundry conditions of piecewise-linear circuit must match,

$$e_i = e_o = \frac{VR_1}{R_1 + R_2},$$

at the clipping threshold.

A negative excursion of e_i will cause a complementary effect with the output saturating when

$$e_i = \frac{-VR_2}{R_1 + R_2}.$$

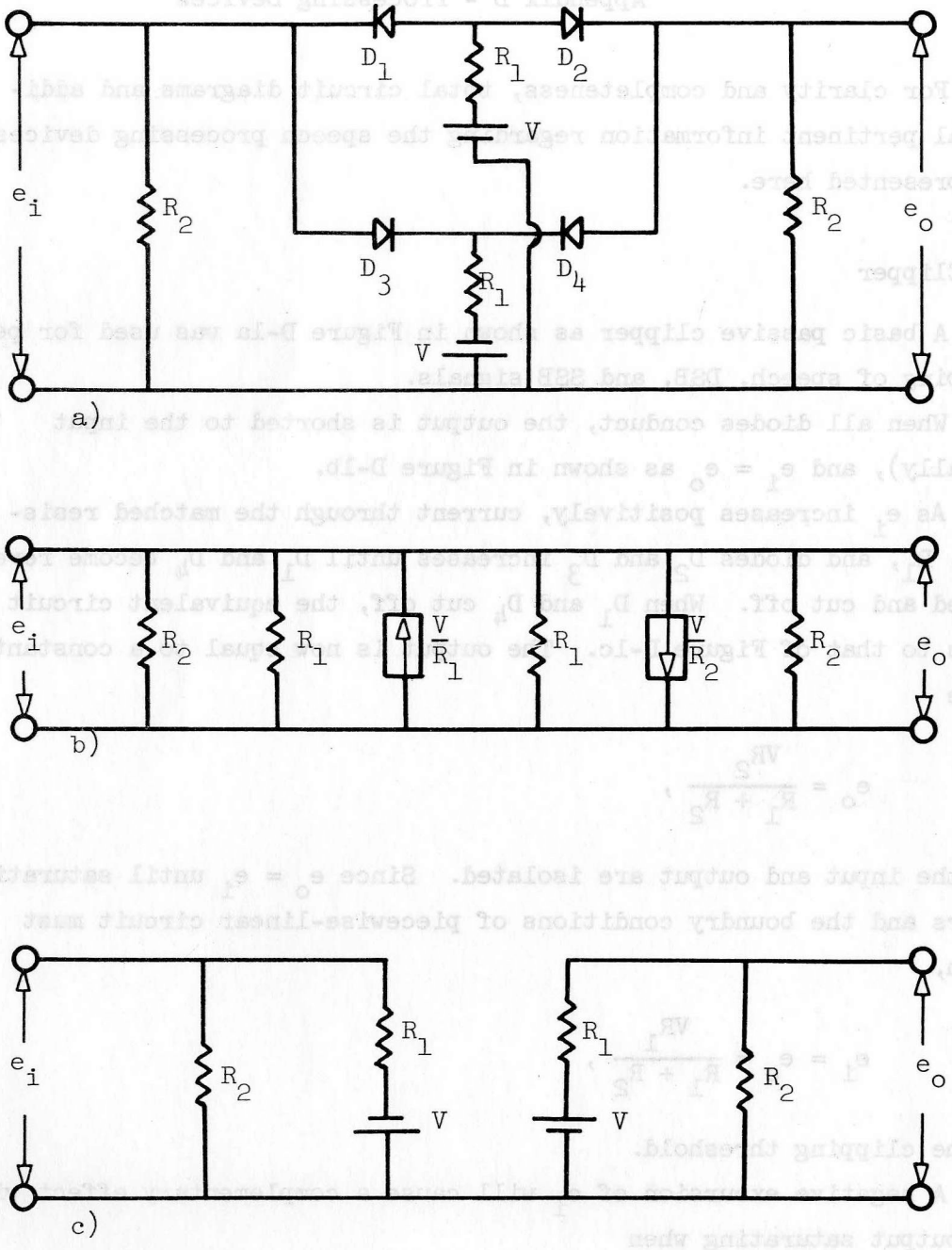


Figure D-1. Theoretical Clipper Considerations a) Basic Clipper Circuit. b) Clipper Equivalent Circuit When all Diodes Conduct. c) Clipper Equivalent Circuit with D_1 and D_4 cut off.

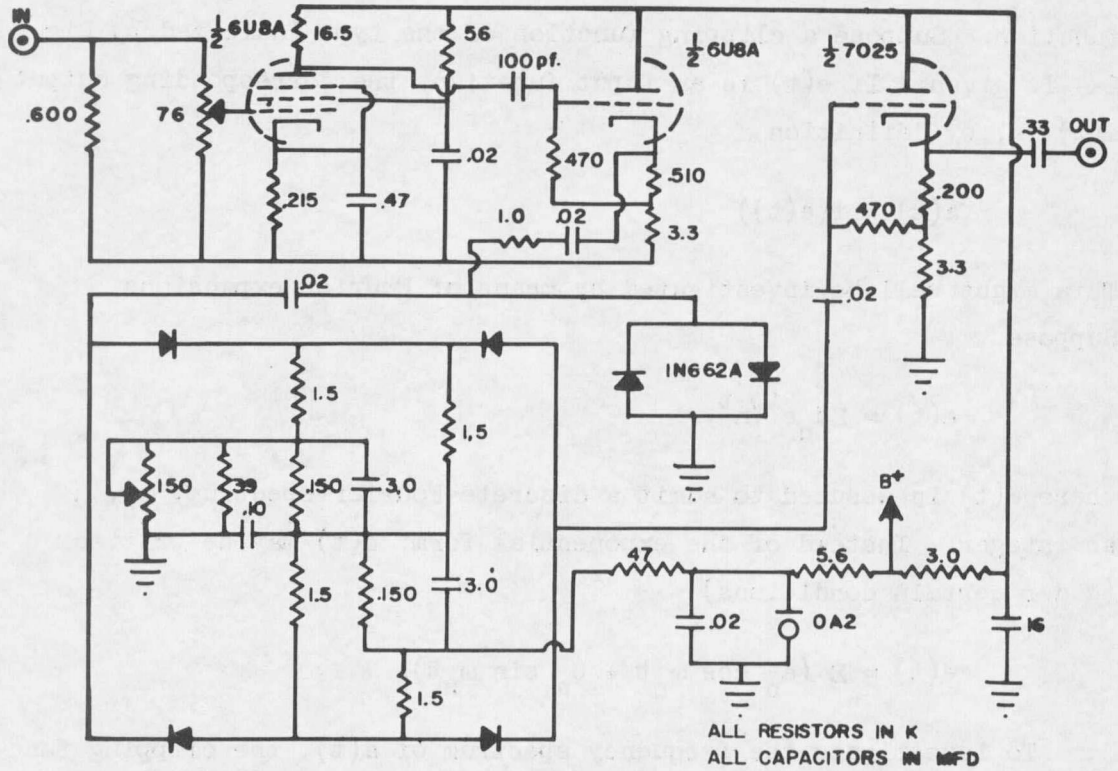


FIGURE D-2 HIGH FREQUENCY CLIPPER

The speech and quadrature function clippers are identical, and are similar in form to the basic clipper of Figure D-1. The high-frequency clippers (Figure D-2) utilize a pair of back-to-back silicon diodes as a pre-clipper to limit signal peak excursions in cases of severe clipping. A buffer amplifier was used to maintain input and output 600-ohm impedance levels.

2. Mathematical Considerations of Clipping*

Mathematically, clipping may be typified in many ways. One is to consider the behavior of Fourier series representing a clipping

*The original formulation of this presentation was made by project consultant, Dr. Hans R. Fischer, Department of Mathematics, Montana State College.

function. Suppose a clipping function of the type indicated by Figure 2-4 is given. If e(t) is an input function, the corresponding output a(t) is, by definition

$$a(t) = f(e(t)).$$

This might well be investigated by means of Fourier expansions. Suppose

$$e(t) = \Sigma d_n e^{i\omega_n t},$$

where e(t) is assumed to admit a discrete Fourier spectrum, ω_n, n an integer. Instead of the exponential form, e(t) may be written (under certain conditions)

$$e(t) = \Sigma\ (a_n \cos \omega_n t + b_n \sin \omega_n t).$$

To investigate the frequency spectrum of a(t), the clipping function f may be approximated by a series of polynomials, $P_n(x)$, uniform in some interval $(-\alpha, \alpha)$, $(\alpha \geq \text{Sup} \mid e(t) \mid)$. This is not a restriction to special cases, since a physical clipper does not yield a truly polygonal transfer characteristic, but rather acts as a sufficiently good polynomial approximation of the form

$$f(x) = \lim_{n \to \infty} P_n(x).$$

If the input signal is assumed to be absolutely covergent in some interval, the output of the approximated clipper,

$$a_n(t) = P_n\ (e(t)),$$

can theoretically be evaluated. First, the powers of the input signal components must be found:

$$(e(t))^k = (\Sigma d_n e^{i\omega_n t})^k = \Sigma\ (\underset{\Sigma n_s}{\Sigma} d_{n_s} e^{i(\Sigma\omega_{n_s})t})$$

In practice, a representation of this form leads to numerical evaluations far too involved and time-consuming. It would be better, therefore to use harmonic analyses to determine the Fourier frequencies and coefficients of $a(t)$; and in general, analyses must be used to determine the expansion of $e(t)$. Furthermore, an expansion for $(e(t))^k$ may be formed only in the case of a finite number of given frequencies, ω_n, since the sum $\Sigma\omega_{n_s}$ must be finite.

On the other hand, suppose the spectrum of the input is given,

$$e(t) = \Sigma d_n e^{i\omega_n t}, \quad \omega_n \neq 0, \quad n \text{ finite.}$$

Also assume a narrow-band signal:

$$|\omega_n - \frac{\Sigma\, \omega_n}{N}| < \epsilon\,,$$

ϵ small for all ω_n, N deciding the total number of non-zero terms. Then an estimate can be made of the number of spectral components in the output which will be in the given narrow-band. This may be illustrated using the "biharmonic case": $N = 2$. Suppose the input is given by:

$$e(t) = a_1 \cos \omega_1 t + a_2 \cos \omega_2 t + b_1 \sin \omega_1 t + b_2 \sin \omega_2 t.$$

Then $a(t) = e^2(t)$, which contains eight components represented trignometrically as:

$$\cos^2 \omega_1 t,\ \cos^2 \omega_2 t,\ \sin^2 \omega_1 t,\ \sin^2 \omega_2 t,\ (\cos \omega_1 t)(\cos \omega_2 t)$$

$$(\cos \omega_1 t)(\sin \omega_2 t),(\sin \omega_1 t)(\sin \omega_2 t),(\sin \omega_1 t)(\cos \omega_2 t)$$

These eight terms may be expressed as:

$$\cos 2\omega_1 t,\ \sin 2\omega_1 t,\ \sin 2\omega_2 t,\ \cos 2\omega_2 t,\ \cos (\omega_1 + \omega_2)$$

$\cos(\omega_1 - \omega_2)t$, $\sin(\omega_1 + \omega_2)t$, $\sin(\omega_1 - \omega_2)t$

and constants.

A second order approximation to the clipping function, therefore, will give the following frequencies in the output:

$$0,\ \omega_1,\ \omega_2,\ \omega_1 + \omega_2,\ \omega_1 - \omega_2,\ 2\omega_1,\ 2\omega_2$$

These may be grouped into three parts:

$$0,\ \omega_1 - \omega_2; \qquad \omega_1,\ \omega_2; \qquad \omega_1 + \omega_2,\ 2\omega_1,\ 2\omega_2$$

$$\text{I} \qquad\qquad \text{II} \qquad\qquad \text{III}$$

The first group consists of frequencies close to zero; the second of frequencies in the given band, and the third of frequencies close together in a new narrow band near $\omega_1 + \omega_2$. That is, only two frequencies will appear in the initial narrow-band, namely ω_1 and ω_2.

A third-order polynomial approximation to the clipping function will yield the following frequencies:

$$0,\ \omega_1,\ \omega_2,\ \omega_1 \pm \omega_2,\ 2\omega_1,\ 2\omega_2$$

$$2\omega_1 \pm \omega_2,\ 2\omega_2 \pm \omega_1,\ 3\omega_1,\ 3\omega_2$$

The groups of frequencies are:

I: $0,\ \omega_1 - \omega_2$

II: $\omega_1,\ \omega_2,\ 2\omega_1 - \omega_2,\ 2\omega_2 - \omega_1$

III: $2\omega_1,\ 2\omega_2,\ \omega_1 + \omega_2$

IV: $2\omega_1 + \omega_2,\ 2\omega_2 + \omega_1,\ 3\omega_1,\ 3\omega_2$

The second group again contains the frequencies which belong to the initial narrow-band. As an example, if $f_1 = 100$ kc and $f_2 = 102$

kc, then group III is of the 200 kc range, group IV is in the 300 kc range etc.

More generally, for an approximation of degree n, group II will consist of frequencies of the form:

$$g\omega_1 - (g-1)\omega_2,\ g\omega_2 - (g-1)\omega_1;\quad 0 < g < n,\quad g \text{ an integer.}$$

The total number of frequencies is 2(n-1). Furthermore, there will be n + 1 groups of terms such that the terms in one group are "close together". The respective ranges are: 0 for group I, the given band for group II; $\omega_1 + \omega_2$ (or $2\omega_1$ or $2\omega_2$) for group III;...; $n\omega_1$ (or $n\omega_2$) for the last group. This is illustrated pictorially in Figure D-3.

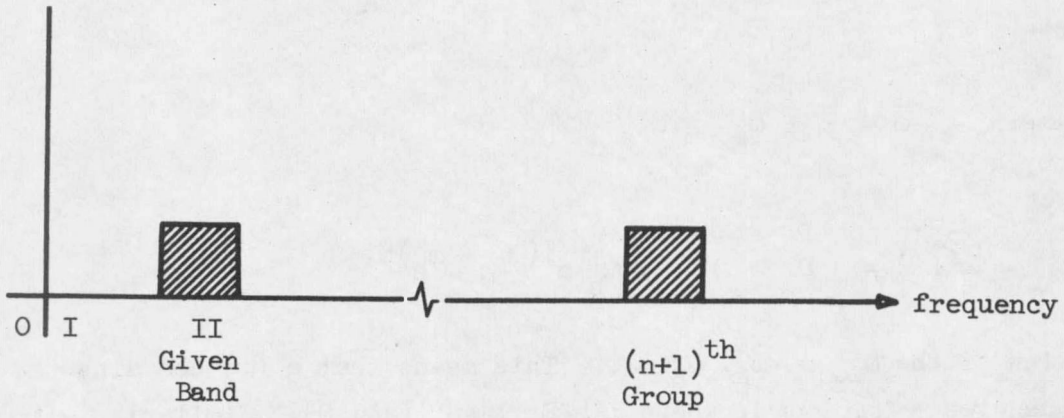


Figure D-3. Spectral Component Groups for an nth-Order Approximation to a Clipping Function

In practice, these relations may be used to determine the order of a sufficient polynomial approximation to the polygonal clipping function: the output frequency components in the given narrow-band (i.e. in Group II) must be found. It will be sufficient to take into consideration only those components, N, which appear with amplitudes greater than a given value. The degree of a sufficient approximation then is $\left[\frac{N}{2}\right] + 1$ or $\left[\frac{N}{2}\right] + 2$ (where $[X]$ denotes greatest integer in X).

In view of the nature of a clipping function, approximation of odd degree should be more convenient.

The foregoing can be simplified by means of the following convention:

$$e(t) = \sum_{-P}^{P} \alpha_g e^{i\omega_g t}$$

where ω_g, $g = 1,\ldots,P$, are the given frequencies which appear in $e(t)$, and where $\omega_{-g} = -\omega_g$, $\omega_o = 0$. The form of $e(t)$ is equivalent to the real form

$$e(t) = a_o + \sum_{1}^{P} (a_g \cos \omega_g t + b_g \sin \omega_g t).$$

Define: $$\bar{\omega} = \frac{1}{P} \sum_{1}^{P} \omega_g$$

Now: $$e^2(t) = \sum_{g=-P}^{P} \sum_{h=-P}^{P} \alpha_g \alpha_h e^{i(\omega_g + \omega_h)t}$$

in view of the $\omega_{-g} = -\omega_g$, $\omega_o = 0$. This means that $e^2(t)$ contains the frequencies $\pm (\omega_h \pm \omega_g)$, where $1 \leq h \leq P$, $1 \leq g \leq P$ (and 0). Similarly, $e^3(t)$ will contain all terms $\pm (\omega_g \pm \omega_h \pm \omega_j)$ where $1 \leq g \leq P$, $1 \leq h \leq P$ and $1 \leq j \leq P$ (and 0). More generally $e^k(t)$ contains those frequencies $\pm (\omega_{g_1} \pm \omega_{g_2} \cdots \omega_{g_k})$, where $1 \leq g_i \leq P$ (and 0). The frequencies which appear close to $\bar{\omega}$ may be written in the form

$$\omega_{g_1} + \omega_{g_2} + \ldots + \omega_{g_s} - \omega_{h_1} - \omega_{h_2} - \ldots - \omega_{h_t}.$$

By re-collecting positive and negative terms, these components are close to one of the original frequencies, ω_n. In a given situation it

is easy to pick out the interesting terms. Let P = 3:

for k = 2 None

for k = 3 The terms: $2\omega_1 - \omega_2$, $2\omega_1 - \omega_3$, $\omega_1 + \omega_2 - \omega_3$

$\omega_1 + \omega_3 - \omega_2$, $\omega_2 + \omega_3 - \omega_1$, $2\omega_2 - \omega_1$,

$2\omega_2 - \omega_3$, $2\omega_3 - \omega_1$, $2\omega_3 - \omega_2$. (Group II)

for k = 4 None

for k = 5 The terms: $3\omega_1 - 2\omega_2$, $3\omega_1 - 2\omega_3$, $3\omega_2 - 2\omega_1$, $3\omega_3 - 2\omega_2$,

$3\omega_3 - 2\omega_1$, $3\omega_2 - 2\omega_3$, etc.

From a table of this nature we could find the number of terms close to $\overline{\omega}$ for an approximation to f(e(t)) of a given degree.

3. Integrator, Differentiator, AVC, Equalizer

The operation of the devices of this section have been described in Chapter II. The circuit diagrams of the differentiator and integrator appear in Figures D-4 and D-5 respectively. The input and output impedances of these devices were 600 ohms, and decoupling was provided in the power supply line to prevent interaction with other circuits.

The circuit diagram of the AVC, shown in Figure D-6, utilized the 6BA6 as a variable gain amplifier. This amplifier is coupled to a cathode-follower whose output impedance is 600 ohms. The control voltage for the variable gain amplifier is derived from a feedback amplifier which drives a full-wave rectifier and filter. The filter has a switched, variable time-constant to provide variable decay times. The amount of control derived is a function of the loop gain control at the input of the 12AU7 amplifier and also the threshold control, abbreviated T.C. on the circuit diagram. The input-output characteristic of

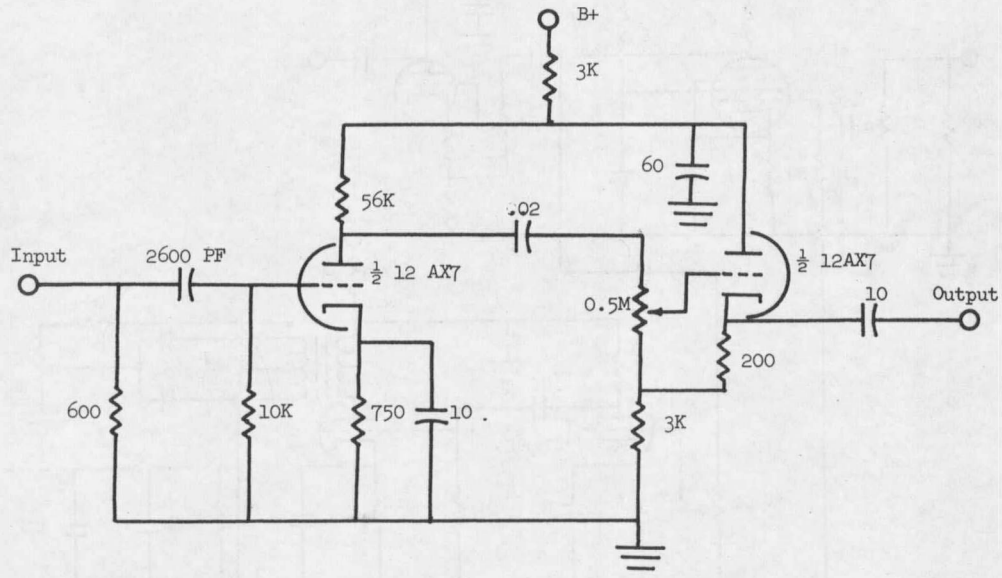


Figure D-4. Circuit Diagram of the Differentiator.

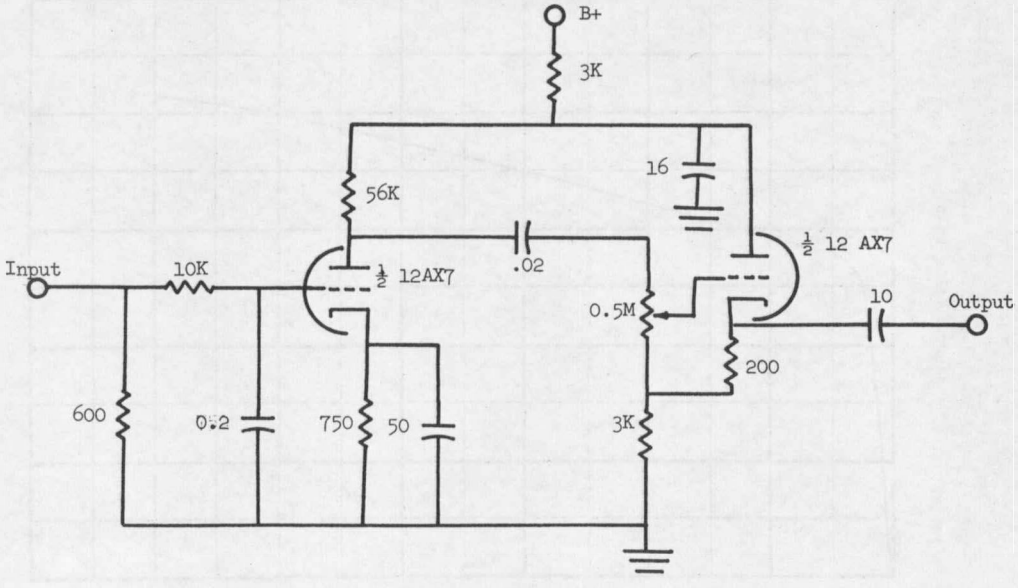


Figure D-5. Circuit Diagram of the Integrator.

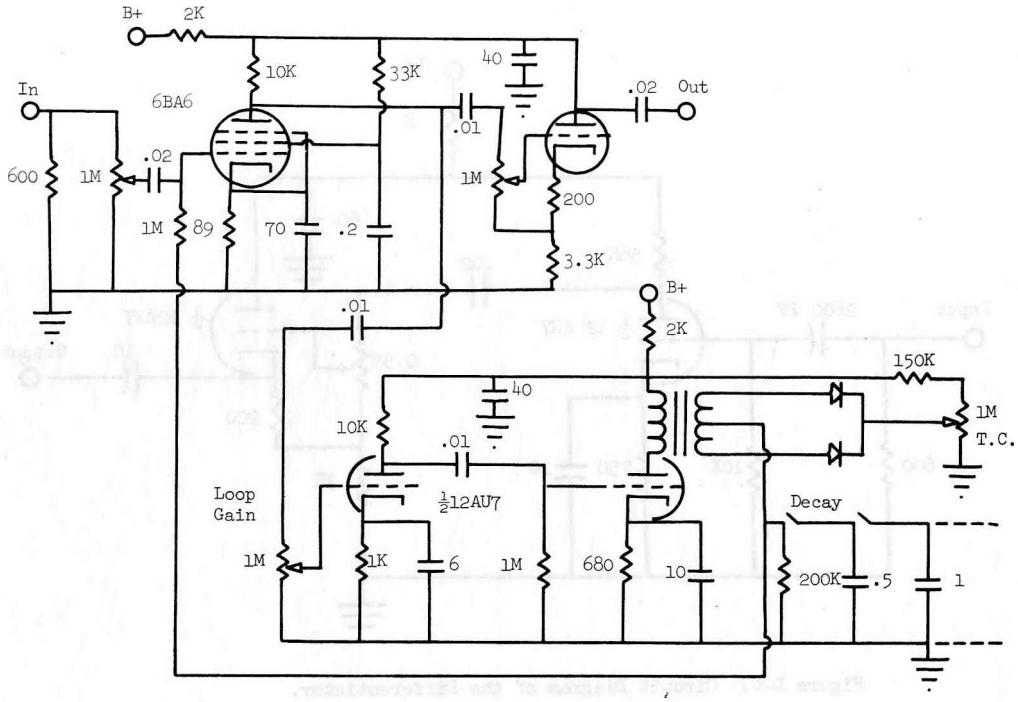


Figure D-6. Circuit Diagram of the AVC.

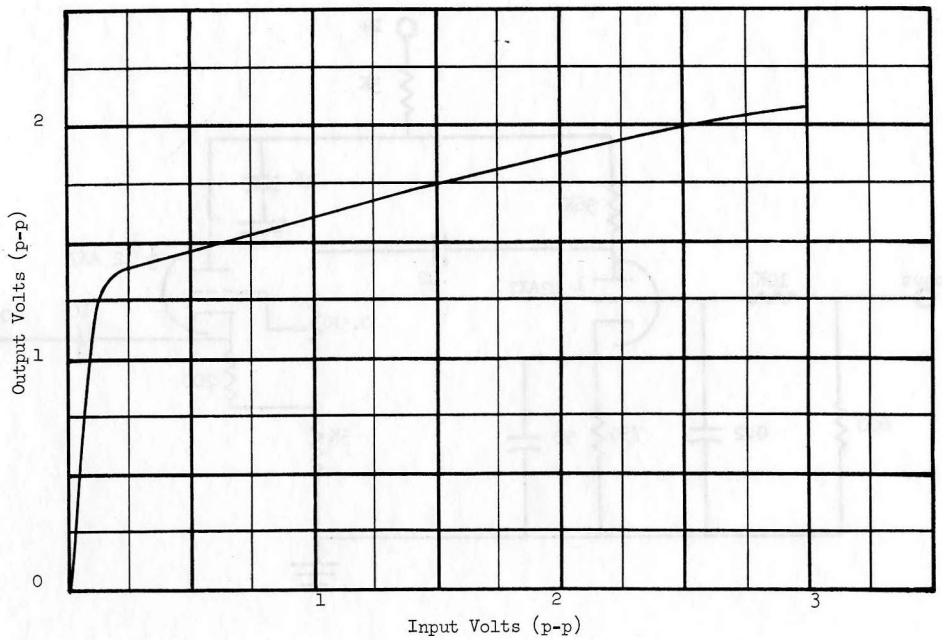


Figure D-7. Input-Output Characteristics of the AVC.

the AVC unit for fixed values of T.C. and loop gain is shown in Figure D-7.

The speech equalizer (Figure D-8) consists of a double integrator and a differentiator with a crossover frequency of 500 cps. A 12AX7 is used to isolate the attenuation networks and to retain system signal level. A variable mixing pad combines the isolation amplifier output signals, which are then reduced to a 600-ohm impedance level by the 6AN8 cathode follower.

A response curve of the speech equalizer is shown in Figure D-9. With a 0.5 volt input signal, the output was computed in db with the 5000 cps point arbitrarily chosen as a reference of 0 db. The quantity M is the ratio (in db) of the output voltage at a frequency, f, to the output voltage at f = 5000 cps.

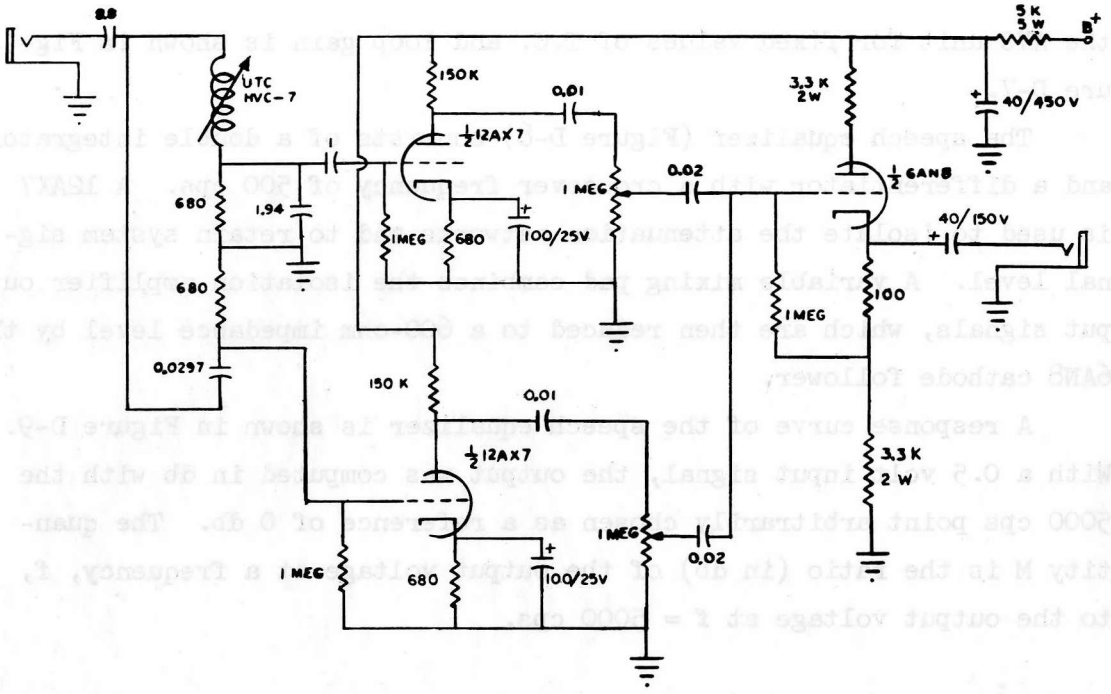


Figure D-8. Circuit Diagram of the Speech Equalizer.

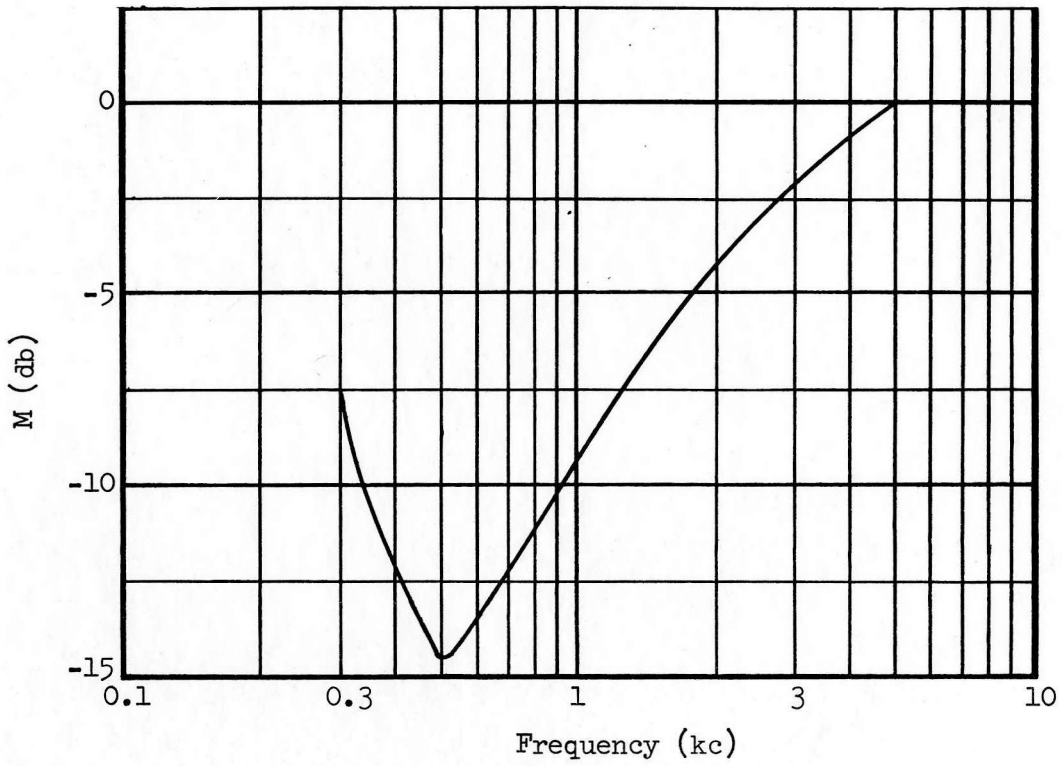


Figure D-9. Speech Equalizer Response Characteristic.

Appendix E - Additional System and Measurement Considerations

1. Earphone Analysis

An initial evaluation of the headsets used for articulation testing was made by means of the General Radio 1551-A Sound Level Meter. These measurements indicated that all earphones were relatively constant in response, one to the other, but also that a peak of approximately 37 db occurred at 2000 cps as compared with the 300 cps response. For these tests the receivers of the earphones were measured individually by holding each receiver to the microphone of a sound level meter.

Since the response shown by the data taken with the sound level meter seemed quite unreasonable, a second evaluation was made using "normal" listeners as a human response meter. Three articulation test team members, with normal hearing as determined by use of the Otometer, took threshold hearing tests by means of the apparatus as shown in Figure E-1. This threshold test was considered to be a summation of system response and listener normal hearing threshold. A normal hearing threshold response curve (published on p. 878 of the ITT Handbook) was then subtracted from the previous threshold response to get a reasonable approximation of system response. Measurement of system response was taken for positions 1 and 2 by this means. These tests indicated that the earphones were within $\pm$ 8 db of flat response from 300 to 5000 cycles.

Use of the data obtained by means of the sound level meter in conjunction with the threshold tests enable a first order evaluation of all earphones used for articulation testing. Such an evaluation for headset #2 is shown in Figure E-2.

During the evaluation of the threshold response data, one subject (C) was suffering from a head cold. This may explain the erratic response given in Figure E-2 for the earphone of position of number 2 at 5000 cycles.

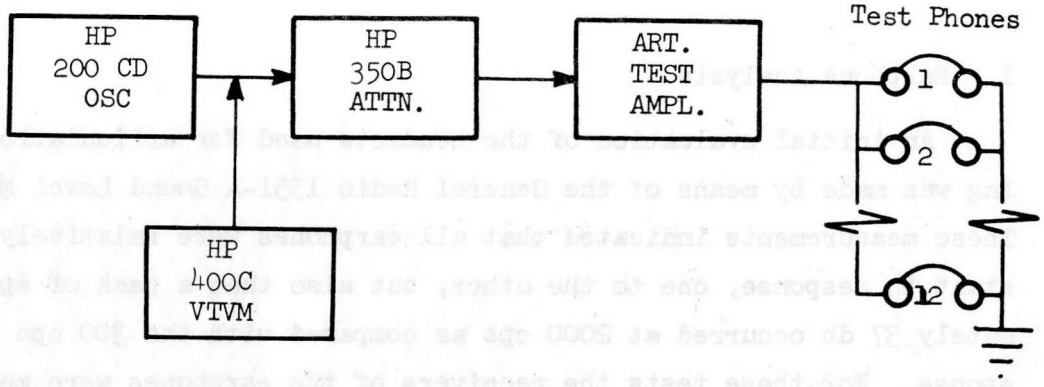


Figure E-1. Equipment Configuration for Threshold Hearing Tests.

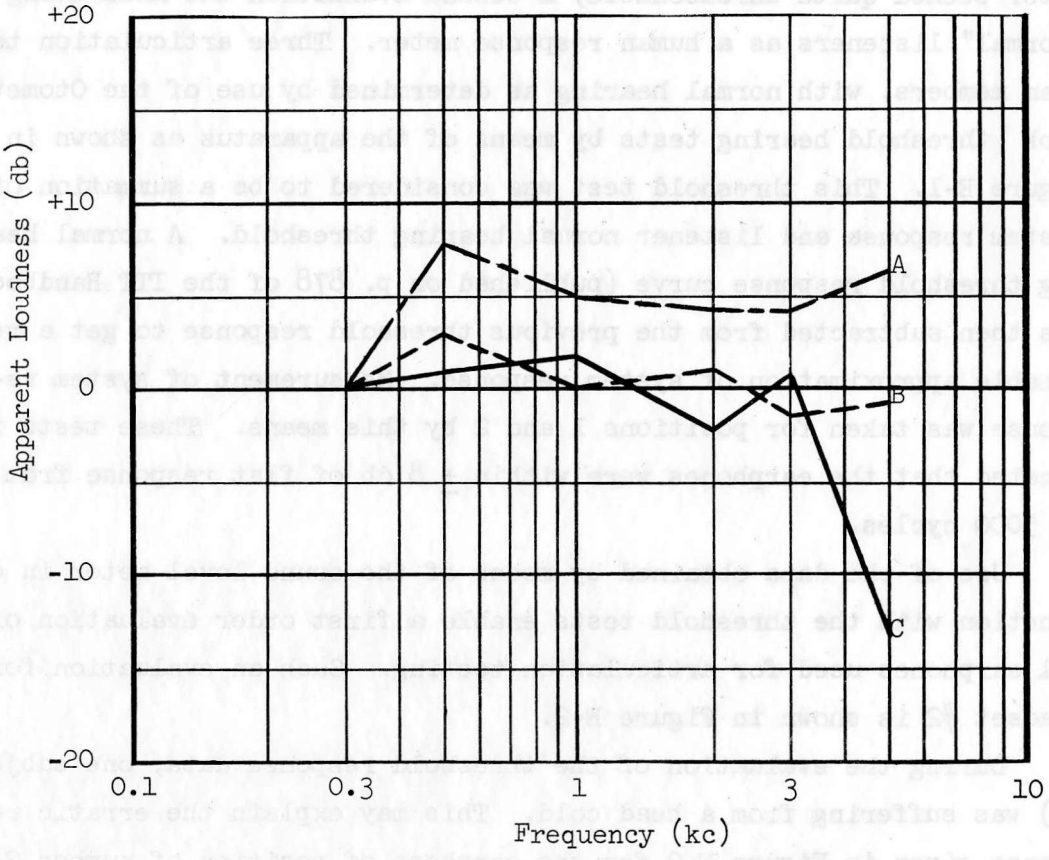


Figure E-2. Response of Earphone, Position #2.

2. True rms vs. Average-reading rms

Voltage measurements of white noise as determined by an rms-calibrated, average-reading meter differ from true rms by approximately 1 db. This may be seen from the following analysis.

The HP 400C Meter gives a reading which is

$$K E_{ave} = K \left[\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |e(t)| dt \right],$$

where K is a calibration constant such that rms readings are given for sinusoidal input signals. If the rms value of a sine wave is normalized to one, $E_{max} = \sqrt{2}$ and the meter input signal may be taken as:

$$e(t) = \sqrt{2} \sin t,$$

which is an odd function having halfwave symmetry. Considering this, the calibration constant for the 400C meter can be determined:

$$K E_{ave} = K \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |e(t)| dt = \frac{K}{2\pi} \int_{0}^{2\pi} |\sqrt{2} \sin t| dt$$

$$= \frac{\sqrt{2}K}{\pi} \int_{0}^{\pi} \sin t \, dt = \frac{\sqrt{2}K}{\pi} (-\cos t)_{0}^{\pi} =$$

$$= \frac{\sqrt{2}K}{\pi} \left[-(-1) + (1) \right] = \frac{2\sqrt{2}K}{\pi} = E_{rms} \text{ (since sine wave).}$$

But $E_{rms} = 1$ for this example, and

$$K = \frac{\pi}{2\sqrt{2}}$$

Now consider the meter reading for white noise (noise having a Gaussian amplitude distribution and constant power spectrum), which

has a mean of zero (no dc bias) and which is normalized to have an rms value of one ($\sigma=1$). The amplitude distribution function is:

$$f(n) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-n^2}{2}\right) \quad (\sigma=1,\ \bar{n}=0)$$

Now,

$$\frac{\pi}{2\sqrt{2}} E_{ave} = \frac{\pi}{2\sqrt{2}} \overline{|n(t)|} =$$

$$\frac{2\pi}{2\sqrt{2}} \int_0^\infty P(n)\, n\, dn$$

but

$$P(n) = f(n) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-n^2}{2}\right).$$

so, $$\frac{\pi}{2\sqrt{2}} E_{ave} = \left(\frac{\pi}{\sqrt{2}}\right) \frac{1}{\sqrt{2\pi}} \int_0^\infty n \exp\left(\frac{-n^2}{2}\right) dn =$$

$$\frac{\pi}{\sqrt{2}\sqrt{2\pi}} (1) = \frac{\sqrt{\pi}}{2}$$

$$db_{error} = -20 \log_{10} \frac{2}{\sqrt{\pi}} = -20(0.052) = -1.04 \text{ db (low)}$$

Thus, the absolute λ conditions as measured with a true-rms meter would be 1.04 db more corruptive than the λ conditions measured with the reading of the HP 400C Meter. For example, $\lambda = 15$ as given in this report would have an absolute value relative to true-rms noise measurement of $\lambda = 14$. This result was verified experimentally.

3. λ Change Due to Clipping

When the PLW of a list is clipped 6 db, other words of that list are not clipped at all. However, channel λ for such unclipped words

is increased.

Let E_p be the normalized peak value in db of the PLW and E_p' be the peak value in db of some other word on the list, such that $E_p' \leq (E_p - 6)$. The clipper outputs for inputs of E_p and E_p' are, respectively (as in Figure 3-9):

$$(E_p - 6) \text{ and } E_p'.$$

The F_m^s outputs are

$$(E_p - 6 + \Delta PV)$$

$$E_p' \pm \delta PV,$$

where $\Delta PV_6 \approx 3$ db (Figure 3-10) and δPV is the peaking (or "flattening") of the unprocessed word due to filter phase characteristics. Note that δPV is associated with the unclipped word; for a typical SSB filter it is generally of the order of $\pm$ 1 db. If β is an additional gain factor (perhaps a linear final amplifier), then the channel signals have relative peak values in db:

$$A_1 = (E_p - 6 + \Delta PV + \beta) = (E_p - 3 + \beta),$$

$$A_2 = E_p' \pm 1 + \beta.$$

Consider now the same words (except unprocessed) in an identical channel:

$$A_3 = E_p + \overline{\beta}\ , \text{ the PLW value}$$

$$A_4 = E_p' + \overline{\beta}\ ,$$

where $\bar{\beta}$ is a different gain factor. If the final amplifier is to have the same peak output, the peak values of the PLW must be the same in each case, so

$$A_1 = A_3$$

$$E_p + \bar{\beta} = E_p - 3 + \beta$$

$$\bar{\beta} = \beta - 3.$$

Then

$$A_2 = E'_p \pm 1 + \beta \text{ (processed)}$$

$$A_4 = E'_p + \bar{\beta} = E'_p + \beta - 3 \text{ (unprocessed)}$$

So, for a given value of E'_p, $A_2 > A_4$; the "processed" word, though unclipped, has a larger value than the same unprocessed word in the channel. Thus, effective channel λ (and intelligibility) of the A_2 word is increased over the no-process case. Note that λ is unchanged for the PLW; however, intelligibility may still be expected to increase over the unprocessed PLW, since the clipping process has reduced the peak factor (by 3 db in this case).

4. Bandwidth and λ.

The increase in severity of a noise condition for fixed λ and decreasing bandwidth may be shown in the following manner.

If the corruptive signal is white noise, by definition the power spectrum is a constant, say k_1 watts/cycle-bandwidth. Now consider a speech signal of a given peak value E_{p_1} and bandwidth f_1. If this signal is mixed with noise from the source with spectral constant k_1,

$$\lambda_1 = 20 \log_{10} \frac{E_{p_1}}{E_{n_1}} .$$

The noise power is:

$$P_1 = k_1 f_1 = G E_{n_1}^2 ,$$

Now if the bandwidth is restricted to a value f_2 (where $f_2 < f_1$) without changing source noise level,

$$P_2 = k_1 f_2 = G E_{n_2}^2 .$$

But if λ is to be constant, E_{n_2} must equal E_{n_1}; so the output of the noise source must be increased to a value such that

$$k_2 f_2 = k_1 f_1 \text{ or } k_2 = k_1 \frac{f_1}{f_2} .$$

Then E_{n_2} is increased to

$$E_{n_2} = E_{n_1} \frac{f_1}{f_2} .$$

A numerical example following Chapter 4 is as follows:

$$f_1 = (5000-300) = 4700 \text{ cps}, \quad f_2 = (3000-300) = 2700 \text{ cps}$$

$$\frac{E_{n_2}}{E_{n_1}} = \frac{f_1}{f_2} = 1.32$$

Or E_{n_2} must be 2.44 db higher than E_{n_1} to maintain λ.

5. Channel λ

In a physical system, a receiver input signal is composed of:

1. An intelligence signal of restricted bandwidth (assuming the transmitter is operating properly).

2. A noise signal, generally of significantly larger bandwidth than the intelligence.

The first stage in the receiver will introduce additional noise, and the bandwidth will be restricted in the IF. Thus, channel λ in an actual system would not have noise and signal bandwidths equal and would not include considerations of front-end noise of a receiver. However, if the receiver is linear through the IF, an "image" channel-λ may be envisioned to correspond to the "channel" used in this investigation, since the definition of λ implies equal signal and noise bandwidths. The fact that a frequency translation is involved between the RF and IF levels is of little consequence; signals at these levels are both narrow-band and the envelope function (hence λ) is invariant with such translation.

Appendix F - Tabular Data

This appendix is devoted to the presentation of tabulated values of N and $\bar{x}$ for the graphical data of Chapter IV. The table numbers presented correspond directly to the Figure numbers of Chapter IV (i.e. Table F-8 is the data for Figure 4-8, etc.). In order to keep the numbering system in order, Tables F-1 and F-2 are trivial since the information contained in these two tables for the examples is contained in Table F-13. All $\bar{x}$ values in these tables have been rounded to the nearest 0.1% intelligibility.

Table F-1. Data for $\mathbb{NP}\,C^{s}_{24}F^{s}_{m}$, $\lambda^{s} = 9$

Process	Degree	λ^{s}(db)	N	$\bar{x}$
$\mathbb{NP}\,C^{s}_{24}F^{s}_{m}$	-	9	59	50.0

Table F-2. Data for $\mathbb{NP}\,C^{s}_{12}F^{s}_{m}$, $\lambda^{s} = 3$

Process	Degree	λ^{s}(db)	N	$\bar{x}$
$\mathbb{NP}\,C^{s}_{12}F^{s}_{m}$	-	3	59	1.0

Table F-3. Data for Initial Control, λ^{a}

Process	Degree	λ^{a}(db)	N	$\bar{x}$
Control (5 kc)	-	0	32	0
Fixed Volume		6	32	8.7
$\mathbb{NP}$		12	28	33.3
		18	32	62.5

Table F-4. Data for Modified Control, λ^{a}

Process	Degree	λ^{a}(db)	N	$\bar{x}$
Control (5 kc)	-	9	22	14.6
Fixed Volume		15	22	50.5
$\mathbb{NP}$		21	22	70.4
		27	22	83.5
Control (5 kc)		9	74	17.9
Var Volume		15	75	50.8
$\mathbb{NP}$		21	75	70.6
		27	76	83.8

Table F-5. Data for E and $T_{f}E$, λ^{a}

Process	Degree	λ^{a}(db)	N	$\bar{x}$
E	-	9	14	33.3
		15	14	65.7
		21	14	71.7
		27	14	83.4
$T_{f}E$	-	9	10	23.0
		15	10	51.4
		21	10	74.0
		27	10	82.4

Table F-6. Data for Audio Clipping Sample, λ^{a}

Process	Degree	λ^{a}(db)	N	$\bar{x}$
$C^{a}_{6}F$	-	9	10	41.8
		15	9	63.6
		21	10	77.4
		27	10	84.4
$C^{a}_{12}F$	-	9	18	54.0
		15	18	75.7
		21	18	86.8
		27	18	91.1
$T_{r}C^{a}_{12}F$	-	9	6	30.3
		15	6	68.7
		21	5	83.8
		27	5	88.8

Table F-7. Data for Audio AVC: Maximum Input Voltage of 0.5 v p-p.

Process	Degree	λ^{a}(db)	N	$\bar{x}$
$V(0.5,X_{1})$	X_{1}=0.1	9	10	19.0
		15	10	50.4
		21	10	82.2
	0.25	9	6	25.7
		15	6	72.0
		21	6	86.3
	0.5	9	40	41.7
		15	40	71.7
		21	40	85.3
	1.0	9	10	39.0
		15	10	69.6
		21	10	83.8
	2.0	9	6	28.0
		15	6	53.0
		21	6	76.0

Table F-8. Data for Audio AVC: Constant Decay-Time of 0.5 sec.

Process	Degree	λ^{a}(db)	N	$\bar{x}$
$V(X_{2},0.5)$	X_{2}=0.3	9	8	42.0
		15	10	69.4
		21	10	80.6
	0.5	9	40	41.7
		15	40	71.7
		21	40	85.0
	1.5	9	8	31.0
		15	7	68.8
		21	7	88.3

Table F-9. Data for Combinations of V and C^{a}_{n}, λ^{a}

Process	Degree	λ^{a}(db)	N	$\bar{x}$
$VC^{a}_{12}F$	-	3	24	28.0
		9	24	63.4
		15	24	79.6
		21	24	88.7
		27	24	93.1
$VC^{a}_{20}F$	-	3	8	30.5
		9	8	65.5
		15	8	87.0
		21	8	88.5

Table F-10. Data for $C_n^a F$ (3 kc Bandwidth), λ^a

Process	Degree	λ^a(db)	N	$\bar{x}$
MP $C_n^a F$	0(Control)	9	36	5.8
		15	76	36.2
		21	36	63.2
	12	3	12	0
		9	12	15.7
		15	12	46.1
		21	12	66.1
	24	3	12	5.2
		9	12	45.5
		15	12	65.0
		21	12	85.0
	36	3	20	12.2
		9	20	55.0
		15	20	81.0
		21	20	89.9
	42	3	12	11.3
		9	12	55.7
		15	12	82.7
		21	12	92.5

Table F-11. Data for $C_n^a F_m^s$, λ^s

Process	Degree	λ^s(db)	N	$\bar{x}$
MP $C_n^a F_m^s$	n=0	9	139	5.8
		15	139	36.2
		21	139	63.2
	6	9	49	13.0
		15	62	43.4
		21	58	64.6
	18	9	49	34.7
		15	62	66.0
		21	48	76.4
	30	9	49	50.2
		15	58	77.4
		21	58	87.6
	42	9	29	43.7
		15	38	72.2
		21	38	86.8

Table F-12. Data for MP $C_n^d F_m^s$, λ^s

Process	Degree	λ^s(db)	N	$\bar{x}$
MP $C_n^d F_m^s$	n=0	9	139	5.8
		15	139	36.2
		21	139	63.2
	12	9	88	17.5
		15	87	52.1
		21	63	66.4
	24	9	86	37.2
		15	87	60.1
		21	63	75.4
	36	9	63	35.4
		15	87	66.0
		21	63	81.1
	48	9	65	30.6
		15	53	67.8
		21	65	86.6

Table F-13. Data for MP $C_n^s F_m^s$, λ^s

Process	Degree	λ^s(db)	N	$\bar{x}$
MP $C_n^s F_m^s$	n=0	9	139	5.8
		15	139	36.2
		21	139	63.2
	12	3	59	1.0
		9	68	31.2
		15	59	61.5
		21	39	76.0
	24	3	59	9.7
		9	59	50.0
		15	67	74.8
		21	38	87.9
	36	3	59	14.9
		9	68	58.8
		15	59	86.2
		21	88	95.0
	48	3	84	19.8
		9	85	65.2
		15	67	91.4
		21	38	94.0

Table F-14. Data for MP $C_{n/2}^s F_m^s C_{n/2}^s F_m^s$ and MP $C_{n/2}^s F_t^s C_{n/2}^s F_m^s$

Process	Degree	λ^s(db)	N	$\bar{x}$
MP $F_m^s C_{n/2}^s F_m^s C_{n/2}^s F_m^s$	n=0	9	38	4.8
		15	38	30.4
	6	9	18	18.2
		15	18	59.1
	12	9	18	33.3
		15	18	71.9
	18	3	18	15.5
		9	36	51.7
		15	36	82.1
		21	18	94.3
	24	9	18	55.5
		15	18	85.5
MP $F_m^s C_{n/2}^s F_t^s C_{n/2}^s F_m^s$	n=0	9	139	5.8
	6	9	20	30.6
	12	9	20	48.1
	18	9	20	61.6
	24	9	20	69.6

Table F-15. Data for $C_n^a F C_n^a F_m^s$, λ^s = 15.

Process	Degree	λ^s(db)	N	$\bar{x}$
MP $C_n^a F C_n^a F_m^s$	n=6	15	18	54.8
	12	15	18	64.4
	18	15	18	68.8
	24	15	18	71.9

Table F-16. Data for Interdependencies: Audio and DSB, λ^s = 15

Process	Degree	λ^s(db)	N	$\bar{x}$
	n_2			
MP $C_6^a C_{n_2}^d F_m^s$	12	15	24	55.0
	18	15	24	59.5
	24	15	24	61.4
	30	15	24	65.3

Table F-17. Data for Interdependencies: Audio and SSB, $\lambda^s = 9$

Process	Degree n_1	Degree n_2	λ^s(db)	N	$\bar{x}$
$\mathbb{P}C^a_{n_1}F^s_m C^s_{n_2}F^s_m$	28.6	18	9	12	69.0
		24	9	12	77.5
		30	9	12	70.0
	39.4	0	9	12	63.0
		6	9	12	75.0
		12	9	12	77.5
		18	9	6	78.0
	44.0	0	9	12	69.0
		6	9	12	80.0
		12	9	12	71.5
	49.8	0	9	12	69.5
		6	9	12	74.0
		12	9	12	66.0

Table F-18. Data for Interdependencies: Audio and SSB, $\lambda^s = 15$

Process	Degree n_1	Degree n_2	λ^s(db)	N	$\bar{x}$
$\mathbb{P}C^a_{n_1}F^s_m C^s_{n_2}F^s_m$	16.6	12	15	20	64.3
		18	15	38	70.7
		24	15	20	76.2
	28.6	12	15	20	79.0
		18	15	38	83.0
		24	15	20	86.0
		30	15	20	87.0
	40.6	6	15	20	88.6
		12	15	20	90.0
		18	15	20	87.6
		24	15	20	79.4

Table F-19. Data for $\mathbb{P}C^a_{n_m}F^s C^s_{36}F^s_m$

Process	Degree	λ^s(db)	N	$\bar{x}$
$\mathbb{P}C^a_{6_m}F^s C^s_{36}F^s_m$	-	3	18	15.1
		9	36	61.4
		15	18	84.9
		21	18	95.0
$\mathbb{P}C^a_{12}F^s_m C^s_{36}F^s_m$	-	3	20	16.8
		9	20	70.1
		15	20	82.3

Table F-20. Data for $T_r C^a_n F^s_m$

Process	Degree	λ^s(db)	N	$\bar{x}$
$T_r C^a_n F^s_m$	n=0	9	56	3.2
		15	56	30.0
		21	18	53.6
	12	9	20	13.7
		15	60	48.6
	24	15	40	65.1
	30	9	20	44.9
		15	20	76.3
	36	15	20	74.9

Table F-21. Data for $T_r C^s_n F^s_m$ and $EC^s_n F^s_m$

Process	Degree	λ^s(db)	N	$\bar{x}$
$T_r C^s_n F^s_m$	n=0	9	56	4.0
		15	56	30.0
	12	9	20	18.9
		15	38	52.6
	24	9	20	47.2
		15	38	75.7
	36	9	20	55.6
		15	20	89.4
$EC^s_n F^s_m$	n=0	9	20	3.2
	12	9	20	18.6
	24	9	20	37.6
	36	9	20	50.9

Table F-22. Data for $T_r C^a_{6_m} F^s C^s_{36} F^s_m$

Process	Degree	λ^s(db)	N	$\bar{x}$
$T_r C^a_{6_m} F^s C^s_{36} F^s_m$	-	3	18	4.4
		9	18	60.9
		15	18	84.5
		21	18	93.5

Table F-23. Data for $\lambda^{27} C^a_{n_m} F^s$, λ^s

Process	Degree	λ^s(db)	N	$\bar{x}$
$\lambda^{27} C^a_{n_m} F^s$	n=0	9	14	4.0
		21	14	62.3
	12	9	9	12.7
		21	9	67.5
	24	9	9	36.0
		21	9	70.0
	36	9	9	35.6
		21	9	70.5
	48	9	9	40.5
		21	9	67.3

Table F-24. Data for $\lambda^{27} C^s_n F^s_m$, λ^s

Process	Degree	λ^s(db)	N	$\bar{x}$
$\lambda^{27} C^s_n F^s_m$	n=0	15	18	34.2
	12	15	18	59.4
	24	15	18	68.6
	36	15	18	70.4
	48	15	18	66.0

Appendix G - Tests of Significance

When graphical data is considered, it is desirable to be able to state whether or not a significant difference exists between any two sample values. For example, in the data of Figure 4-13, does a significant difference exist between the scores for n = 24, λ = 3 and n = 24, λ = 9 (the points marked 1 and 2)? This question was answered by using a one-tailed Mann-Whitney U test. The U test is one of the most powerful of the nonparametric tests:

> "if the U test is applied to data which might properly be analyzed by the most powerful parametric test, the t test, its power-efficiency approaches $3/\pi = 95.5$ per cent as N increases" (Siegel, 1956).

Let us assume that all individual differences and systematic errors have been reduced to a minimum by the selection and training of the test team and by the various precautions used in the recording procedure (see Chapter 3). Then any two sets of scores for different n or λ of Figure 4-13 may be considered independent samples. The null hypothesis is, then, that population 1 and population 2 have the same distribution; the alternate hypothesis, H_1, is that the sample values of 2 are (stochastically) larger than the sample values of 1. Let the significance level be, arbitrarily, $\alpha = .001$; the rejection region then consists of all values of Z which are so extreme that their associated probability under H_o is equal to or less than $\alpha = .001$.

Let the samples be of size n_1 and n_2, then

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}, \qquad \text{(G-1)}$$

where

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1, \qquad \text{(G-2)}$$

or

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2, \qquad (G\text{-}2)$$

whichever is smaller.

R_1 = sum of ranks assigned to the group n_1

R_2 = sum of ranks assigned to the group n_2

Formula (G-1) does not take into account the effect of tied scores; however, unless the number of a given tied score approaches $\frac{n_1 + n_2}{2}$, satisfactory results are obtained by using the form (G-1).

A numerical example for C_n^s, $n = 24$, $\lambda = 3, 9$ is given below. In Figure 4-13, these points are labeled 1 and 2, respectively, and the sample scores for these points are:

1. 26,14,12,6,20,24,12,6,2,18,14,14,18,22,20,26,12,8,2,18,6,0, 20,4,8,20,2,28,16,14,10,2,8,6,4,8,6,0,10,4,8,0,0,0,0,4,0,2, 4,4,2,16,4,12,14,2,14,16. ($\mathbb{N} = 58$)
2. 56,56,62,54,64,62,54,58,50,56,60,66,60,50,62,80,56,54,60,76, 44,38,72,32,42,54,40,58,52,40,38,36,54,28,34,38,42,52,32,32, 46,44,70,26,34,44,44,46,34,54,42,66,46,52,42,52,62,52,40. ($\mathbb{N} = 59$)

A ranking of these scores yields the table of the following page. The sum of ranks for each sample is:

$$R_1 = 1713.5, \; R_2 = 5189.5.$$

Now, $n_1 + n_2 = 58 + 59 = 117$ and a necessary condition is:

$$R_1 + R_2 = \sum_{n=1}^{117} n$$

which is satisfied for this example.

The value of U can now be calculated by using (G-2):

$$U = 58(59) + 1/2\,(59)(59 + 1) - 5189.5 = 2.5$$

Intelligibility Score	Ranked Score of: 1	Ties	2	Ties
0	4	7		0
2	11	7		0
4	18	7		0
6	24	5		0
8	29	5		0
10	32.5	2		0
12	35.5	4		0
14	40.5	6		0
16	45	3		0
18	48	3		0
20	51.5	4		0
22	54	1		0
24	55	1		0
26	57	2	57	1
28	59.5	1	59.5	1
30		0		0
32		0	62	3
34		0	65	3
36		0	67	1
38		0	69	3
40		0	72	3
42		0	75.5	4
44		0	79.5	4
46		0	83	3
48		0		0
50		0	85.5	2
52		0	89	5
54		0	94.5	6
56		0	99.5	4
58		0	102.5	2
60		0	105	3
62		0	108.5	4
64		0	111	1
66		0	112.5	2
68		0		0
70		0	114	1
72		0	115	1
74		0		0
76		0	116	1
78		0		0
80		0	117	1

$R_1 = 1713.5$ $R_2 = 5189.5$

Then, by (G-1):

$$Z_{12} = \frac{2.5 - 1711}{\frac{58(59)(118)}{12}} = \frac{-1708.5}{183.4} = -9.31$$

or $|Z_{12}| = 9.31$

Reference to Table A (Page 247, Siegel, 1956) reveals that $|Z| \geq 3.43$ has a one-tailed probability under H_o of $p \leq .0003$. Since this p is smaller than $\alpha = 0.001$, the decision is to reject H_o in favor of H_1. We thus conclude that the scores of point 2 are (stochastically) higher than the scores of point 1. For the point-to-point data of Figure 4-13, most of the calculated values of $|Z_{ij}|$ were greater than 3.43; consequently, the values of $|Z_{ij}|$ may be used as an indication of relative significance for several i,j sample-point pairs:

Sample point i,j	Z_{ij}	p
1,2	9.32	<.00003
2,3	8.65	<.00003
3,4	6.13	<.00003
5,6	7.35	<.00003
6,2	6.79	<.00003
2,7	3.73	<.00003
7,8	2,68	0.0037
8,9	1.42	0.0778

Note that the null hypothesis, H_o, may not be rejected at the $\alpha = .001$ level for the point pairs 7,8 or 8,9. Further, H_o may not be rejected for 8,9 until α is greater than 0.0778.

MONTANA STATE UNIVERSITY LIBRARIES
3 1762 10011210 9

D378
Sh98
cop.2

Shyne, N. A.
Speech-signal processing &
applications to single-sideban

	NAME AND ADDRESS
2-3-63	Warren Leyte 404 Fairfield
NOV 2 1964	Orville Hughes 1206 DEER St.
4-11-65	James L. Dorsey 15 Faculty MSU Bldg
6-30-66	Dan Marsh, ERL
7-1-66	Dan Marsh ERL
9-26-66	Dan Mc
Oct. 1-66	Dan
5-6-67	RUO
7-30-74	
OCT 5	
OC	

D378
Sh98
cop. 2