6767

NATIONAL LIBRARY
OTTAWA

CANADA

BIBLIOTHÈQUE NATIONALE
OTTAWA

NAME OF AUTHOR.....Walter J. Watts.....

TITLE OF THESIS...A verbal and electro-acoustic Analysis of Transient Emotional States.....

UNIVERSITY.....U. of A.....

DEGREE FOR WHICH THESIS WAS PRESENTED.....Ph.D.....

YEAR THIS DEGREE GRANTED.....1970.....

Permission is hereby granted to THE NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(Signed)..W. J. Watts.....

PERMANENT ADDRESS:

...115 Burpee St.
...Fredericton, N. B.
.....

DATED. Sept. 10.....19 70

NL-91 (10-68)

THE UNIVERSITY OF ALBERTA

A VERBAL AND ELECTROACOUSTIC ANALYSIS OF TRANSIENT EMOTIONAL STATES

by

© WALTER JULIUS WATTS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1970

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, "A Verbal and Electroacoustic Analysis of Transient Emotional States" submitted by Walter Julius Watts in partial fulfillment of the requirement for the degree of Doctor of Philosophy.

Supervisor

External Examiner

Date Sept. 1972

ABSTRACT

The purpose of this study was to investigate the feasibility of using electroacoustic analyses of different frequency bandwidths in order to detect changes in transient emotional states during continuous speech production.

The subjects in this study comprised two groups of eighteen males each. One, called the NON-CLINICAL, was selected from the under-graduate population of the University of Alberta. The other group, referred to as CLINICAL in this study, was drawn from the workers of the Sheltered Workshop in Edmonton.

The study required subjects to read prepared neutral (NSS) as well as prepared emotional (ESS) speech passages into a high-fidelity tape-recorder. In addition to these passages subjects were asked to produce approximately five minutes of spontaneous speech in which they described some dramatic event in their lives. Using an anxiety scale which is applicable to verbal samples, two anxiety-loaded (ASS) and two anxiety-free (AfSS) samples were obtained from each subject's dramatic monologue.

Intensity reading in millivolts were obtained for the NSSs, ESSs, ASSs, and AfSSs produced by each subject in these frequency bands: 80 - 250 Hz, 200 - 900 Hz, and 80 - 6300 Hz. Mean Speech Pressure (MSP) readings between

pairs of passages (NSS - ESS, AfSS - ASS, ESS - ASS, and NSS - AfSS) were computed for each subject's speech samples in all three frequency bands. These data entered the statistical analyses.

Three primary hypotheses together with their four respective sub-hypotheses were generated and tested for each group of subjects within each frequency band. The primary hypotheses and their sub-hypotheses, identical for each group of subjects and frequency band, read as follows: Changes in the level of arousal during continuous speech production will result in concomitant changes in vocal intensity within this frequency range, while parity of levels of arousal remains unaffected. The four sub-hypotheses were stated this way: (a) There is a difference in MSP between NSSs and ESSs. (b) There is a difference in MSP between AfSSs and ASSs. (c) There is NO difference in MSP between ESSs and ASSs, and (d) There is NO difference in MSP between NSSs and AfSS.

Tests on the primary hypotheses resulted in the unanimous acceptance of all four sub-hypotheses covering the 200 - 800 Hz band, for both groups, thus substantiating the primary hypothesis. The hypotheses covering the 80 - 250 Hz band were partly accepted for the NON-CLINICAL

group, but had to be rejected for the CLINICAL one. The primary hypotheses covering the 80 - 6300 Hz bandwidth were rejected for both groups of subjects. From the above results, the conclusion can be drawn that the 200 - 800 Hz bandwidth reflected changes in level of arousal or parity between levels of arousal most clearly.

The availability of data suggested a comparison of speech intensities between the two groups of subjects involved. A set of secondary hypotheses were tested which compared the Mean Speech Intensity of the NON-CLINICAL group to that of the CLINICAL one for each of the three bands. Without exception, the Mean Speech Intensity of the CLINICAL group was significantly lower than that of the NON-CLINICAL group.

ACKNOWLEDGEMENTS

The writer wishes to express his deep appreciation to those people who made the completion of this thesis possible.

The author is indebted to Dr. Harvey W. Zingle, thesis advisor, for his advice and criticism during the course of the study.

The writer wishes to gratefully acknowledge the untiring assistance of Dr. Kenneth H. Gough on whose knowledge and understanding the writer depended so often, particularly in regards to the technical execution of the research.

Appreciation is also due to Dr. Paul C. Sartoris for his refreshing approach in dispelling 'thesis blues' and for the giving of encouragements at times like these. The author extends his appreciation to all those students who participated in this study for giving their free time, but he is particularly grateful to those workers from the Sheltered Workshop who trusted him in sharing some of their personal life experiences with him. Association with these men was the most unforgetable part of this research.

The author would like to express his appreciation to Shirlee Heft for her competence and perseverance while typing this material.

Und zum Schluss denn, Dir, Mady, gilt mein aufrichtigster Dank und ein frohes 'Glück Auf'.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

THE THESIS PROBLEM AND ITS BACKGROUND

THE THESIS PROBLEM AND ITS BACKGROUND

I Introduction

Neither novel nor controversial is the viewpoint that an interpersonal exchange communicates more than a stenographic transcription of the same conversation. Words are not the only property of speech which serve to convey information about the feelings of the speaker. It is well known that the manner of speaking affects the listener's perception of the speaker's feeling state.

Language behavior is only one part of the total interactional process which consists of the encoding and decoding of exchanged messages through a multi-channel expressive system. About the relative importance of three of these channels of the expressive system, Mehrabian and Wiener (1967) state:

> ...that the combined effect of simultaneous verbal, vocal and facial attitude communication is a weighted sum of their independent effects - with the coefficients of 0.07, 0.38, and 0.55, respectively (p. 252).

The differential weighting of verbal and vocal aspects of speech is illustrated in the commonly quoted statement that 'it is more important how you say it than what you say'. A large part of the emotional message is expressed to a great extent through the vocal content. Diehl (1960) expressed this notion very well when he wrote:

> All the parts of the vocal apparatus, from the diaphragm to the lips are intimately related through

> the autonomic nervous system with the function of the internal organs. The neural effects of emotional disturbance in the internal organ are unconsciously transmitted to the various parts of the vocal mechanism. Manifestations of muscular tension throughout the vocal mechanism will thus conform to the nature of the emotional state, with the specific quality and intensity of each emotion experienced having their counterpart in the human speech (p. 175).

For the listener speech has essentially two components which are the intentionality of the verbal message and the manner of speaking. The latter reveals important aspects of the speaker's relative stable personality characteristics and his more transitory emotional states. Acoustically, speech can be considered consisting of three dimensions, namely intensity, frequency, and duration. Previous research (Fletcher, 1949; Kramer, 1963) has suggested that the emotional component is carried by the vocal aspects of speech which is related to the frequency range employed in expressing emotions. While low frequencies apparently carry much of the speech power, high frequencies, which are responsible for the intelligibility of conceptual material, carry very little.

While the majority of studies investigating emotionality or arousal in voice using continuous speech have relied on the reports of listener-judges, very few studies have acoustically analyzed speech beyond the single monosyllabic response for this purpose. The present study is an investigation of the practicality of using electroacoustic analyses of different frequency bands for the detection of transient emotional states in continuous speech.

II Problem

The number of studies related to the topic of speech and expression of emotionality or arousal through this channel of communications are quite numerous. In the predominant number of cases, researchers have relied on the subjective reports of listener-judges. Starkweather (1961) stated that this form of judging produced higher interjudge correlations than the external criteria seem to warrant. With the improvement of electroacoustic hardware over the last two decades, several researchers have profitably exploited new avenues in the study of the relationships between speech and emotions.

Most of the studies employing electroacoustic analyses consider speech falling along two continuums: intensity, or the loudness dimension; and frequency, the high-low dimension. Intensity readings in the above studies are expressed either in pressure or power units, and pitch, in cycles per second or Hertz (Hz).

For some time, psychologically oriented researchers have investigated the effects of variations of frequency bands on the intelligibility of speech as well as on the expression and perception of emotions. Research on the perception of emotions or arousal has frequently made use of content-free speech samples. In order to render a sample of speech content-free, that is, to remove frequencies above a certain range, the voice sample is

filtered using bandpass filters which permit only designated frequencies to pass. Voices with frequencies over 500 Hz removed are reduced to a kind of a mumble as though heard through a wall. One still hears indications of pitch, rate, and loudness, though probably a good deal of what is usually called voice quality is lost together with the high frequencies and the content. With respect to the conveying of emotions, Soskin and Kauffman (1961, p. 80) sum up the findings very well when they state that "voice sounds alone, independently of semantic components of vocal messages, carry important clues to the emotional state of the speaker".

A step further removed from listeners' judgment is the study by Friedhoff, Alpert, and Kurtzberg (1962) who induced a mild transitory emotional state by having their Ss tell a lie. This was accomplished by having each S respond to a series of stimuli containing the numbers from one to ten, with the response 'No'. Prior to the experiment, Ss were told to choose one of the ten numbers, which they were told not to relate to others. Subjects were then presented with lighted number stimuli to which they responded with 'No'. All responses were tape-recorded. The signal from the tape recorder was passed through a rectifier-filter circuit, amplified and recorded on a polygraph (Offner Type R, Dynograph), with the deflection of the polygraph pen indicating the sound pressure level of each S's voice. The frequency band ranging from 80 - 6300 Hz was used in this

study. A masking noise was used which was piped into the S's ears via earphones. The masking noise was at about 80 dB re 0.0002 dyne/cm^2 in order to prevent Ss from self-monitoring and adjusting their voices while replying to the stimuli. It also became apparent that each S adjusted to the emotional stimulus individually either by increasing or decreasing the intensity of his voice; however, any given S tended to change his voice in the same direction on all trials.

A study by the same team (Alpert, Kurtzberg, and Friedhoff, 1963) using very much the same experimental outline, analysed the 'No' responses in the 100 - 250 Hz and the 80 - 6300 Hz frequency band. No masking noise was used. The results showed significant differences between three pre-ES (emotional stimuli) and three post-ES trials for the low frequency band with probabilities ranging from less than 0.01 to 0.80. The highest probabilities were found following the ES presentation, indicating a carry-over effect in the immediately succeeding trials.

Another study using monosyllabic speech responses is the one by Rubenstein (1966) who investigated the intensity and frequency dimensions of vocal responses to limited stress. The 'Yes' or 'No' responses of a group of young women in answer to questions about their emotional state during a short stay in a sound-proof room were recorded before and after treatment. The treatment consisted of 10 minutes of

isolation in this darkened room. An analysis of variance showed a significant decrease in intensity after isolation. There was no change in fundamental vocal frequency. A group of listener-judges was not able to determine consistently whether the responses were made before or after treatment, i.e. the stress situation in the suddenly darkened booth. A second experiment which consisted of the administration of electro-shock showed significant changes in intensity as well as in fundamental vocal frequency. No attempt was made in these studies to isolate intensity changes in different frequency bands.

A study aimed in a slightly different direction is the one by Holmgren (1967). This researcher investigated the physical and psychological correlates of speaker recognition. He obtained results in a factor analysis using listener-judged and actual physical voice measures which indicated that five factors accounted for more than 90 per cent of the total variance. Holmgren speculates that listeners make their judgment about speakers on the basis of (1) the mean of the frequencies in the 70 - 145 Hz range, (2) the mean amplitude and variance of the amplitude of the frequencies in the 70 - 3000 range, and (3) rate of speech. It is of particular interest to the present study to find that listeners are able to make vocal-perceptual judgments on changes in those frequency ranges studied, which are included in the bandwidths investigated in the present study.

It becomes apparent from the preceding studies of Alpert and associates, (1962, 1963), Rubenstein (1966), and Holmgren (1967), as well as those on content-free speech that certain frequency ranges contain sufficient clues to the emotional information conveyed.

The present study is an attempt to investigate the feasibility of analyzing transient emotional sequences occurring in longer samples of speech using an electroacoustic analysis of voice intensities in three frequency bands. Two of the frequency bands studied - the narrow band between 100 - 250 Hz, and the broad band between 80 - 6300 Hz - have been employed before in investigations. In addition to these two bands a third one is analyzed. This band comprises the frequencies between 200 - 800 Hz and encompasses the range of the first formant frequencies of all vowel phonemes. Studies using content-free speech have almost consistently included all or a portion of this range in their analyses, but have offered little justification for their selection of bandwidths. The present study included the 200 - 800 Hz bandwidth with the assumption that vocal intensity changes would be reflected in the vowel production and in particular that vowel resonance region which contains the greatest speech power (Fletcher, 1949; French and Steinberg, 1947).

In addition to the electroacoustic analysis of transient emotional sequences in continuous speech, a verbal measure of transient emotional states is employed in the present

investigation, too. This measure is an anxiety scale developed by Gleser, Gottschalk, and Springer (1961) and derives its scores from transcribed material of continuous speech samples. The rationale for the inclusion of this scale in this study is its reported ability to select emotional speech samples from transcribed material.

This study consists of the following steps:

(1) tape-recording of the S's reading voice while reading six selected passages which carry no obvious emotional content as judged by a body of undergraduate students.

(2) tape-recording of the S's reading voice while reading six selected passages containing emotional content as judged by a body of students.

(3) tape-recording of up to five minutes of speech after requesting the subject to relate about some dramatic, personal events in his life.

(4) transcribing and scoring of verbal material obtained under step (3) using the GLESER anxiety scale.

(5) electroacoustic analyses of neutral and emotional reading passages as well as of selected passages obtained under step (3). The analyses use the following frequency bandwidths: 80 - 250 Hz, 200 - 800 Hz, and 80 - 6300 Hz.

CHAPTER II

REVIEW OF RELATED RESEARCH

REVIEW OF RELATED RESEARCH

The material which is reviewed on the following pages has been organized under five major sub-headings according to the area and type of research reported. The sub-headings are: Review of Earlier Interview Content-Analyses; Voice and Speaker; Voice and Emotion; Stereotypes in Voice Judgments and, Measuring of Emotions.

I Review of Earlier Interview Content-Analyses

In a review of the literature of studies concerned with the content-analysis of psychotherapeutic interviews, Auld and Murray (1955) state that early studies have suffered basically from three hindrances: (1) the basic data of therapy are transient and accessible to the therapist only; (2) conclusions arising out of investigations were matters of impressions and opinions because of the lack of objective verbal material analysis and, (3) the data could not be fitted into a suitable theoretical frame work. The use of sound recording techniques have helped to overcome the first two drawbacks to a considerable extent, and psychoanalysis has laid the groundwork for the conquering of the third in that it provided a theory of personality and psychopathology.

Auld and Murray classify their studies according to three classes: (a) studies aimed at developing measures of analysis - or methodological studies; (b) descriptive studies and, (c) studies of cause-effect relationships.

Representative of the methodological studies are the systems and measures developed by Snyder, Curran, Bales, and Dollard and Mowrer. Snyder (1945, 1947) classified the therapist's responses according to the technique employed: restating content, clarifying feelings, interpreting, structuring, leading, accepting, etc. The client's responses are categorized into: problems, insight, planning, and simple responses containing questions, answers, utterings of disagreement, etc. Curran (1945) obtained measures of 'insight' which he defined as the common ground between two different problem areas. In addition to this work, Curran was concerned with the classification of the problem. Curran wanted to know whether the issue was one of hostility, discouragement, feelings of inferiority and so on. Bales (1950), developed the Interaction Process Analysis which involves a series of steps like getting information, making decisions, and carrying out actions with one participant asking for information, another giving it, and the third one judging the opinion offered. Dollard and Mowrer (1947) postulated that responses are the result of drives which have been reinforced through drive reduction. By considering each word either as a drive, reward or neutral, a Discomfort-Relief Quotient (DRQ) is obtained with words suggesting discomfort like suffering, tension, and pain in the numerator and relief and discomfort words in the denominator.

Some of the descriptive studies have employed the DRQ. Assum and Levy (1948) in their study of one case found a significant drop in DRQ which they interpreted as success. Cofer and Chance (1950) computed the DRQ for each interview hour of their five published cases which were judged to be successful by therapists. All five cases showed a drop in DRQ. Concerning the value of the DRQ as a correlational measure of 'success', Auld and Murray (1955) stated that unless an adequate measure of 'success' is available, a drop in DRQ is not of much value for assessment purposes.

Among the theoretically guided studies, Lasswell's pioneering work is worth mentioning. Lasswell (1935) hypothesized that 'conscious affect' material consists of references to the therapist while 'unconscious tension' is indicated by slow speech, pauses, and interruptions. He found correlations between his measure of tension and physiological indices like the GSR. Snyder (1945), working with various types of responses by the therapist, found that statements revealing insight and the discussion of plans resulting out of insight were more likely to follow or emerge out of nondirective than directive responses by the interviewer. Dittman (1952) found that slightly interpretative responses produced more progress responses in the client than purely reflective ones did.

The content-analysis systems described in the preceding section are not adequate to the task of marking out the main

variables in therapy. Measures of the content of clients' and therapists' utterances have to be supplemented by measures of other non-verbal responses of client and therapist. The significant communality among the studies reported in this section is the global attempt these researchers used to pry the important from the total verbal message. Their tools of investigation were crude and their results reflect this.

II Voice and Speaker

Kramer (1963, p. 408) states that "A person's changing emotional state and relatively stable personal charcteristics may be judged from nonverbal properties of his voice". Allport and Cantril (1934) and Cantril and Allport (1935) reported a series of 14 experiments which involved 24 speakers and 600 judges. Judges were asked to match voices with twelve features of personality covering the overall appearance of the speaker, both in person and from photographs. Some success was achieved, but no characteristic was always revealed correctly, and in general, judges agreed with each other in excess of their accuracy. A series of studies by Fay and Middleton investigated judgments of a number of personality characteristics from voice transmitted over a public address system. Among those characteristics judged with success were Spranger's personality types (Fay and Middleton, 1939), but they had little success with such attributes as sociability (Fay and Middleton, 1941), and

introversion-extroversion (Fay and Middleton, 1942), using the Bernreuter (1931) Inventory.

When comparing test scores and listener judgments on certain traits with particular voice characteristics, Moore (1939) found that subjects with a "breathy" tone of voice tended to be lower in dominance and higher in introversion. Mallory and Miller (1958) reported that Bernreuter scores on introversion were negatively related to loudness, low pitch, and resonance in the voice.

Starkweather (1955, 1956), utilizing a technique developed by Fletcher (1949), French and Steinberg (1947), and Licklider and Miller (1951), which consists of removing those frequencies of the speech spectrum which are necessary for the carrying of verbal content, contrasted speakers with high and low scores on a personality test (Harris, 1953) which distinguishes hypertensives from normals. His hypothesis that the hypertensive syndrome group would show greater incongruence between verbal and vocal aspects of speech was not confirmed. In another study Starkweather (1956b), using the above filtering technique, found that judges were able to separate submissive from aggressive subjects both from normal speech as well as content-free or filtered samples. He found that filtered samples had a slight advantage in judgment. Judges who were to make the distinction between the two groups on the basis of transcripts only were not able to perform the task.

The identification of speakers, using expressive differences which are likely related to personality variables, has been discussed by Kersta (1962a, 1962b) and Smith (1962). Hargreaves and Starkweather (1963) in a related study using a digital speech spectrometer system, spectrum analyzer and other interconnected equipment like a scanner which receives the output and feeds it into a digital voltmeter which in turn takes a two-digit reading from each of the 18 channels used and prints the data on teletype paper, were able to predict the speaker with an accuracy score ranging from 60 to 100 per cent, with a total across the whole group of subjects of 90 per cent accuracy.

In a study centered around Eysenck's description of extroversion-introversion, Ramsey (1968) using three groups of subjects, all Dutch, found that introverts employed longer periods of silence between utterances which were taken as a sign of higher cognitive activities thus confirming Eysenck's theory.

In order to determine the number and nature of basic ways in which voices are perceived and distinguished by listener-judges, Voiers (1964) had 32 judges rate 16 voice samples on a 49 item semantic differential scale which were subjected to an analysis of variance and later factor-analyzed to obtain values for the dimensionality of speaker effect, listener effect, and listener-speaker interaction effect. With respect to the speakers, four factors labeled

clarity, roughness, magnitude, and animation accounted for 88% of the variance in mean ratings given to speakers. The listeners' dimensions were clarity, conspicuity, masculinity, belligerence, and tautness. These accounted for 57% of the common-factor variance in constant errors. Five dimensions, labeled pleasantness, roughness, magnitude, unnamed, and hardness, accounted for 33% of the common-factor variance of the interaction between speakers and listeners.

Addington (1968) studying the relationships of selected vocal characteristics to personality perception, recorded specific passages spoken by two male and two female speakers simulating seven different voice qualities: tense, breathy, thin, flat, throaty, nasal, and orotund; three variations of speaking rate, namely normal, fast, and slow; and normal, higher than normal, and lower than normal pitch variations. Employing several groups of judges who described the vocal characteristics on a nine-point-equal-interval scale for each of the three dimensions, Addinton found the following personality decriptions to emerge from a factor analysis: lanky-dumpy; hearty-glum; potent-impotent; and soft-hearted - hard-hearted, for males; and gregarious-antisocial; aggressive-unresisting; urbane-coarse, hardy-fragile; and appealing-disagreeable, for females. In general, the factors so isolated suggest that male personality was judged in terms of physical and emotional power, while female

personality was perceived in terms of social faculties. With respect to speech rate, male and female speakers were considered more animated and extroverted when it increased. Pitch increase in males was considered as more dynamic, feminine and showing aesthetic inclinations; in females it was perceived as more dynamic and extroverted.

It has been recognized by psychiatrists and clinicians (Moses, 1954; Ostwald, 1961 and 1965; Soskin, 1953; and Sullivan, 1954) and others that the nonverbal or vocal aspect of speech can be utilized for the diagnosis of psychopathology and in its treatment in therapy. Sullivan writes:

> ...the psychiatric interview is primarily a matter of vocal communication, and it would be a quite serious error to presume that the communication is primarily verbal. The sound-accompaniments suggest what is to be made of the verbal propositions stated (1954, p. 5, 7).

Moses (1960) stated that voice is the expression of 'physiological manifestations' and the emotions expressed through it are part and parcel of it. In a comparative study of schizophrenic children with normal ones, Goldfarb, Braunstein, and Lorge (1956) found that the former group of children was much less effective in conveying and perceiving mood or emotion vocally than the latter group.

Ostwald has attempted to provide objective measures of psychopathological conditions detectable from vocal aspects of speech and of the influence of treatment on

these measures. By analyzing the recorded voice of patients in such a way that the sound is denoted in form of a graph with frequencies in Hz in the abscissa and intensity in half-octave band levels in decibles re 0.0002 microbar as the ordinate, Ostwald (1960a, 1960b) was able to obtain voice prints or curves for each patient. These voice prints can be superimposed for each patient so that fluctuations in voice over units of time can be studied. The half-octave band measurements of intensity show acoustic energy levels to be higher in certain frequency bands than in others, with normally four distinguishable peaks. Ostwald related the results of his work with patients to four distinct categories of human sound which he described as follows: (1) 'sharp' voice, suggesting excitation and complaint; (2) 'flat' voice, produced by depressed and obsessional people who present an irritating but indistinct acoustic facade; (3) 'hollow' voice, reflecting depression, stupor and/or organic brain disease and/or damage and, (4) 'robust' voice, suggesting vigor, found mostly in extroverted, aggressive, and confident people. (See Figure I, page 18)

A report on three cases, a 64-year old man with diagnosed manic-depressive psychosis, a 16-year old girl during an acute schizophrenic state, and a 32-year old woman in an acute psychoneurotic hypochondriacal state, presents statistically significant changes in loudness in

the acoustic spectrum centering around 500 Hz, and a more robust type of voice after psychiatric treatment (Ostwald, 1961). Ostwald (1965) made a frequency analysis of the speech of different types of patients showing that the speech of depressives is flat, low-pitched, and monotonous, whereas that of manics has a robust, resonant quality.

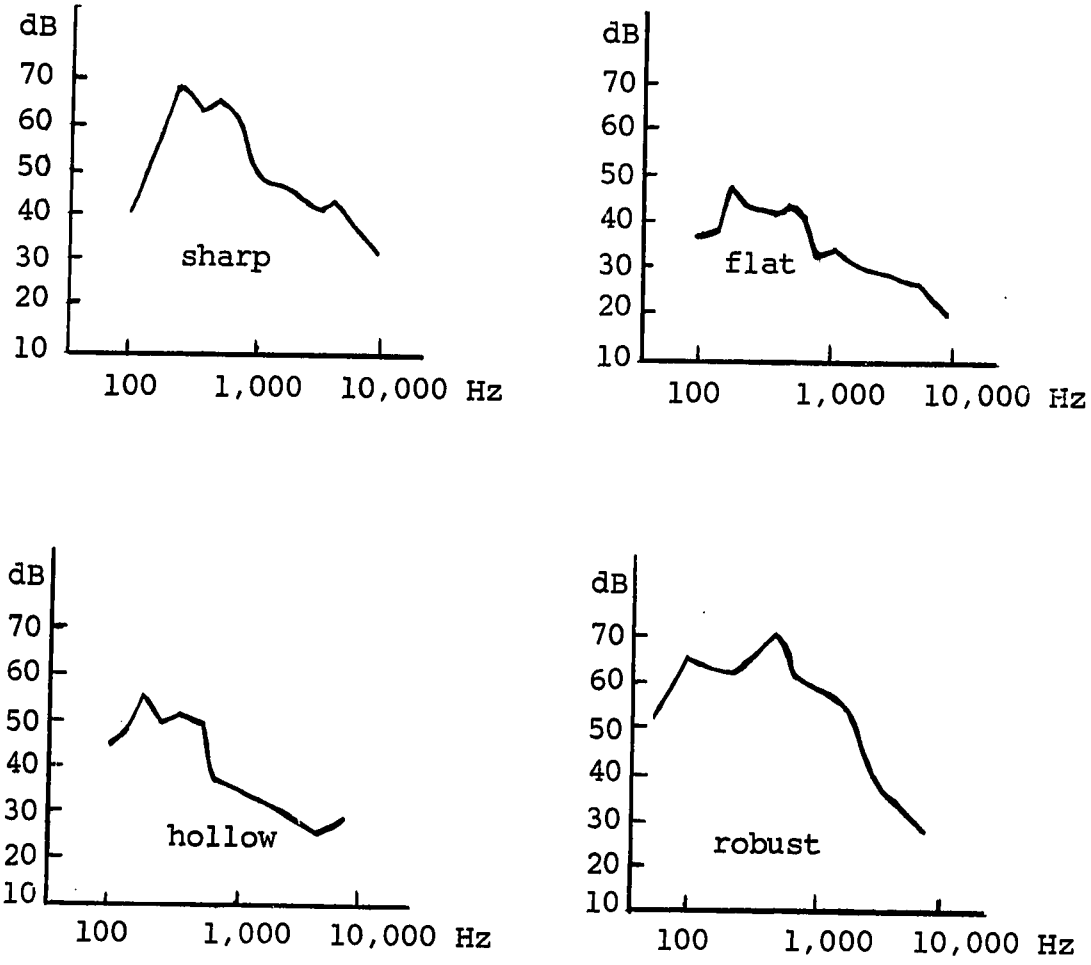


Fig. I; Adapted from: Ostwald, P. F. Soundmaking. Springfield, Ill.; Charles C Thomas, Publisher, 1963.

Paraphonia is the term used to describe the disagreement between vocal aspects of speech and the content of the conversation. Spoerri (1966) in a study of the speaking voice of 350 schizophrenics, who were noted for their speech irregularities, mimic and gestural behavior, found when comparing electroacoustic readings of pitch, duration and intensity of three normals who uttered the same passage as spoken by the patients, that the profiles produced by the schizophrenics were flat and showed continuous sound. The author, who draws a distinction between speech which is primarily informative, and the speaking voice in which the expressive values stand in the foreground, concluded that in the case of his patients the expressive values are preeminent over the informative and communicative ones, producing a strong discrepancy indicative of the sender's monologue. The work of Goldfarb, et al., (1956), as well as that of Moskowitz (1952) and Markel, Meisels, and Houck (1964) suggests that schizophrenics have distinct voice qualities which are distinguishable from those of non-schizophrenics.

Hargreaves, Starkweather, and Blacker (1965) explored a method for the measuring of voice quality rather than speech content of depressive patients. Using a spectrometer, the recorded sound patterns were punched on punch cards representing the voice spectrum which is then fed into a computer. The computer produces a spectrum (average)

for each consecutive five seconds of speech. In a comparison between mood ratings by two clinicians and voice spectra obtained from depressives, accurate predictions could be made for 25 of the 32 patients using voice spectra.

In 1962, Weintraub and Aronson published an article in which they described the application of twelve verbal measures to mannerisms and verbal operations which are indicative of ego defenses in terms of psychoanalytic theory. The isolated verbal measures are: quantity of speech, length of pauses and silences, rate of speech, retractors (although, except, nevertheless), explanations (due to, as a result of), direct references, expressions of feeling, and evaluators. Applying these measures to the ten minute verbal samples of each of ten women and five men diagnosed for impulsive behavior and to a control group of 23 Armed Service men, the researchers (Weintraub and Aronson, 1964) found their hypothesis confirmed which stated that impulsives react to stress by manipulating - appealing for help - authority figures in order to remove the distressing cause of their discomfort, thus producing high 'direct reference' scores as well as high 'retractor' scores reflecting their desire to undo. Another study (Weintraub and Aronson, 1965) with delusional patients also confirmed their hypothesis which had predicted higher scores on 'direct references', 'negators', 'explanations',

and 'evaluators'. In a comparative study between depressives and a control group, Weintraub and Aronson (1967) found significant differences for the twelve measures with the exception of 'qualification', 'retraction', and 'explanation'. Depressives produced fewer words, had a lower rate of speech, and fewer 'non-personal references', but they exceeded the control group in all other remaining measures. In a different study (Aronson and Weintraub, 1967) verbal productivity was taken as an operational measure of energy output of depressive patients. Improvement of their condition was defined in terms of change on the MMPI D-scale. Those patients who improved during treatment showed a return to normal verbal productivity from either hypo- or hyper-productivity which was idiosyncratic for the individual patient.

The preceding section is indicative of the need for an instrument or technique which can aid in the objective quantification of psychologic processes and conflicts as expressed in the clinical interview or in social interaction. Several systems have been proposed depending largely on the special orientation of the respective researchers. The lack of a strong, unifying theory is apparent, though despite its lack or vagueness, some interesting progress has been made.

III Voice and Emotion

Verbal communication has been described by Soskin (1953) in terms of two channels. That part of speech which carries the articulated sound patterns called words, Soskin referred to as the 'semantic' channel; and the nonverbal, or vocal aspect, the 'affective' channel. The existence of a two-channel theory had been postulated by others. Of interest here are the studies by Fletcher (1953), French and Steinberg (1947), and Licklider and Miller (1951). In studying emotional states as portrayed through voice, researchers have used several methods. While some used meaningless content like numbers or 'Ah' sounds, others have kept the content constant by using sentences of neutral quality.

Skinner (1935) recorded the 'Ah' sound of subjects who had read some lines of emotional literature and had listened to selected pieces of music which were supposed to put them either into a happy or a sad state of mind. The 'Ahs' of happiness showed greater force and higher pitch. Thompson and Bradway (1950) found a significant correlation between two sets of statements on the 'affective interchange' as judged by the two psychologists who had acted out a therapeutic interview using numbers as means of communication. These numbers were spoken with the normal inflections reflecting the emotions of client and therapist.

Davitz and Davitz (1959), having speakers recite the alphabet with varying expressions, reported that listener-judges could correctly identify the expressions far beyond chance expectations. Their success was not uniform over all the emotions portrayed. Anger, for example, was identified correctly 65 per cent of the time, while love and pride, less than 25 per cent of the time. Fear was commonly identified as nervousness; love was misjudged as sadness, while pride was taken as satisfaction.

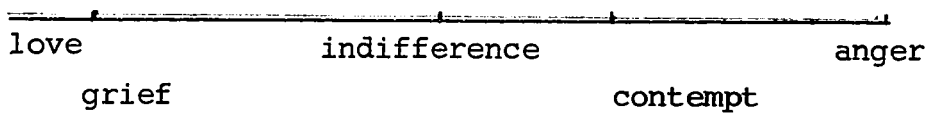
In the case of constant content the same words are used to express different emotions. Using experienced student-actors and listener-judges who heard only a set of sentences that was common to all five passages portraying different emotions, measurable pitch differences (Fairbanks and Pronovost, 1939) and differences in duration of phrases (Fairbanks and Hoaglin, 1941) were found among the different emotions, using average measures from the different readings.

Another promising approach in the study of vocally portrayed emotions is the filtering technique which has been mentioned briefly before. In filtered speech the verbal content is eliminated by using low pass-band filters which pass frequencies below a certain range only while holding back those which carry verbal content. Soskin and Kauffman (1961) had two groups of judges rate the

emotional content of 15 speech samples, with one group listening to filtered speech in which frequencies over 450 Hz had been filtered out. The two sets of ratings showed high correlational agreement. Kauffman (1954) stated at the completion of an analysis of various filtered and unfiltered speech sample judgments that there is a strong tendency for the verbal or semantic channel to function in a manipulative sense, and the vocal or affective channel in an expressive one. In a less artificial situation, Starkweather (1956b) had twelve clinical psychologists judge three speech samples each of Senators Welch and McCarthy during the 1954 Army-McCarthy hearings. During the first two presentations, judges were to rate the filtered speech samples, the frequencies of which were severely attenuated above 300 Hz, according to the 'amount of emotion expressed' and the pleasantness, in the samples. Judgments on unfiltered speech were also made. The ratings showed high interjudge agreement.

In addition to using filtered and unfiltered samples of speech, Kramer (1964a) employed a foreign language, Japanese, in his study on vocally transmitted emotions. Four sentences of varying lengths were incorporated into passages reflecting anger, contempt, grief, indifference, and love. The filtered passage in the English language had frequencies above 400 Hz attenuated. Judgments of both

filtered and unfiltered English passages were compared showing 46 per cent agreement between the two presentations, a result which differed significantly from chance. Investigating individual differences among judges in their ability to rate filtered as well as unfiltered speech samples, Kramer found no significant variation. Judges were less accurate in their rating of emotions conveyed in foreign speech. While anger, grief, and indifference were recognized fairly clearly, contempt and love were predominantly judged as indifference. Dawes and Kramer (1966) subjected the data of the preceding study to a proximity analysis in which the expressed emotions were "represented in a space consisting of a single dimension" (p. 574). The degree of similarity, graphically demonstrated below, of two emotions is expressed in terms of the correlation coefficient, r, between the two emotions. Their difference (d) on the uni-dimensional space is a function of $d^2 = 2(1-r)$ using Coombs (1964) criterion which states that $d(a,b) < d(a,c)$ in the space representing emotions a, b, and c, provided emotion b, is more similar to emotion a than is emotion c.



A great number of findings in the present review suggest that vocal variables are as sensitive indicators of transient emotions in man as are content or lexical variables. In a study concerned with estimating the amount of anxiety from vocal as well as content variables of speech, Gottschalk and Frank (1967) found support for a theory of 'redundancy' of lexical and vocal aspects of speech rather than an 'additive' theory, suggesting that anxiety ratings can be as validly obtained from lexical as from vocal factors. The authors based their conclusion on the insignificant correlational differences between two rating procedures. An anxiety rating utilizing the Overall-Gorham Scale (Overall and Gorham, 1962) obtained from typescript plus sound recording, was compared with anxiety scores derived from a scale developed by Gleser, Gottschalk, and Springer (1961), resulting in a product-moment correlation coefficient of $r = 0.86$; a similar comparison without sound recording produced a correlation of $r = 0.78$. Another team of researchers (Mehrabian and Wiener, 1967) also investigating two-channel verbal-vocal decoding, concluded however, that in the case of incongruence of message between the two channels, normal addresses subordinated the verbal component to the vocal one. A study by Mehrabian and Ferris (1967) supports the view of an 'additive' theory.

As far as the study of speech is concerned, some attempts have been made, as can be taken from the above review of the literature, to analyze the communicative functions of variations in tones of voice, speed of utterances, problems of intonation, and so forth. Many such studies adopted a method analogous to the one most widely used in the study of facial expressions, in which subjects are presented with samples of speech varied in a number of ways. It has been found that people are able to judge what emotion is being expressed through variations of this sort. There has been little research on these variations as they actually occur in interaction, however, and much useful work would be possible.

IV Stereotypes in Voice Judgments

In the great majority of studies concerned with the investigation of relationships between sound of voice and personality and/or emotions, judges have listened to groups of speakers attempting to match voice and personality or voice and emotions. Starkweather (1961, p. 65) in a review of the literature concerned with this problem was left "pessimistic concerning the utility of assessing such traits from nonverbal stimuli". Echoing the sentiment of earlier writers (Sanford, 1942; and Licklider and Miller, 1951), Starkweather ascribed the quite frequent findings that listeners showed higher interjudge correlations than

the external criteria warrant to the existence of stereotyped voices. Since actors have been used rather frequently in portraying various emotions or types of personality, Cowan's admonition that actors emphasize cultural stereotypes of emotional expressions, is well taken (Cowan, 1936).

Kramer (1964b) in a very interesting paper suggested that "this interjudge agreement is not without validity, and that the role of seeking correlations with external criteria has not been fully understood in such studies" (p. 247). Kramer, after citing Campbell (1960, p. 248) on the description of trait validity, stated that listener judgments are as valid a measure of a trait as are the test scores which have been used for the external criteria.

While several studies have dealt with the differences among speakers, Kramer feels that studies investigating personality differences and other relevant variables among listener-judges and their influence on perception, have been ignored. Four areas deserving attention are:

(1) listeners' motivational need structure and its influence on judgment of others.

(2) age differences among judges, particularly children vs. adults.

(3) differential acuity to sound and the perception of personality and emotions in others.

(4) cultural-linguistic variables or the variation of non-verbal cues expressed by different language groups.

Shapiro (1968, p. 181) relating about the lack of unanimous agreement among judges stated that "essentially, the analysis of ratings of human behavior, in which the raters' behavior is considered the variable under study, is an analysis of the perceiver in person perception".

V Measuring of Emotions

One of the general problems involved in the quantification of affect is the formulation of a satisfactory working definition of affect. Theoreticians and experimentalists have approached this problem in different ways, and these differences have sometimes led to varying conclusions.

Mahl (1956), working with transcripts of psychotherapeutic interviews, established a number of verbal measures based largely on formal criteria which were later validated by comparison with the content of verbatim material. The most promising among his methods are the speech disturbance measures: 'Ah' sounds; sentence change; repetition of words; stutter; omission of words or parts of them; incomplete sentences; tongue slips, and intruding incoherent sounds. Later studies, Cook (1969), Kasl and Mahl (1958, 1965) and Mahl (1958) as well as Paivio (1965) suggested however, that the 'Ah' sound is not a useful index of transient anxiety. The other speech disturbances (SD) became known as 'Non-Ahs' and are expressed in ratio form with 'Non-Ahs' over the total number of words in the

sample under study.

Paivio (1965) and Kasl and Mahl (1965) found no relation between permanent anxiety and Non-Ah SDs, while Zimbardo, Mahl, and Barnard (1963) reported a significant interaction between permanent anxiety, in the high transient anxiety condition, and subjects with low test anxiety, in the low transient anxiety condition, had higher rates of Non-Ah SDs. The latter result is strange; the low permanent anxiety-low transient anxiety group is the group that should show the least sign of anxiety.

Studies more concerned with transient emotional states, notably Dibner (1958), Kasl and Mahl (1965), Zimbardo, et al. (1963), Krause and Pilisuk (1961), and Pope and Siegman (1962) have found that transient anxiety leads to increases in Non-Ah SDs. Several of these studies present physiological data or self-report data, showing that the manipulative action was effective. In a very recent study, Cook (1969), using two recognized measures of permanent anxiety, the Taylor Manifest Anxiety Scale and the McReynolds Assimilations Scale (AS) (McReynolds and Acker, 1966) obtained no significant relation between the Non-Ah SDs and any of the two scales, suggesting the the SD measure is a function of transient anxiety only. Speech rate variations have been considered by some researchers as indicators of emotional states. The results of Cook's study, which incorporated speech rate (SR) as one of the

two dependent measures of anxiety, reflect the previous inconclusive findings of earlier investigations. Feldstein, Brenner, Jaffe (1963) and Kanfer (1960) reported an increase in SR when subjects related anxiety-producing situations; while Boomer and Dittman (1964) and Siegman and Pope (1965) found a decrease or non-affection in SR respectively.

A number of workers have attempted to investigate the psychotherapeutic process using a variety of research tools in order to assess the client's emotional state. The investigation of Strupp (1960) employed a five dimensional system for the analysis of therapist behavior and characteristics. Rating scales tapped the overall behavior of the therapist, the degree of inference of warmth in his communication, the degree of channeling or manipulating of the client, and use of therapist's or client's frame of reference. Other studies on the therapeutic interaction have been reported by Jaffe (1957, 1958) and his associates Fink, Jaffe, and Kahn (1960) and Jaffe, Fink, and Kahn (1960) who developed the 'dyadic' system in which the verbal material of client and therapist is considered as emanating from one person. A tool in their analysis is the Type-Token Ration (TTR) which is the ratio of different words within the unit under study over the total number of words employed. The TTR is a rather sensitive measure reflecting changes in defense maneuvres and at critical points during

therapy. Jaffe, using this measure in his analysis of psychiatric interviews, concluded that extremes of stereotype and diversity are indicative of grossly pathological communication and that successful treatment should be accompanied by an increase in variability.

Among the studies in the nonquantitative analysis of interview content are the attempts by linguists like Pittinger and Smith (1957), Pittinger, Hockett, and Danehy (1960), and Trager (1958) to apply linguistic techniques to the analysis of emotional expressions in speech. Using speech samples of three minutes duration, Dittman and Wynne (1961) coded these according to linguistic as well as paralinguistic phenomena. Linguistic characteristics so coded were juncture, the dividing point in speech separating clauses; stress, the pattern of increase and decrease of loudness within clauses and pitch, the rising and falling of the fundamental frequency. Paralinguistic phenomena, so called by Trager (1958, p. 4) are voice quality, consisting of tempo, rhythm, rasp and resonance in the voice; voice set, the physiological characteristics of the speaker and vocalization, composed of vocal characteristics like laughing, crying, voice breaking, etc.; vocal segregates like 'ahs', 'um-hmms'; and vocal qualifiers like unusual changes in intensity, pitch, and variations in duration. Dittman and Wynne stated however, at the end of

their report, that while the more traditional linguistic categories such as juncture, stress and pitch would be reliably coded, they bore no relationship to the emotional states of the speaker; the paralinguistic categories, however, were expected to be more emotionally relevant but could not be reliably coded.

Some of the most provocative research in this category has been done by Gottschalk and his co-workers who have made a number of important contributions to the verbal behavior literature. They have concentrated their effort on the development of scales of anxiety, hostility, and schizophrenia based on psychodynamic trends as revealed by thematic material in the client's speech. Gleser, Gottschalk, and Springer (1961) developed an anxiety scale applicable to verbal samples arising out of interviews providing data on transient emotional changes and sequences during the therapy hour. Gottschalk, Gleser, and Springer (1963) presented a hostility scale applicable to verbal samples with subscores for overt, covert, and ambivalent hostility. A scale attempting to measure different schizophrenic disorganizations has been developed by Gottschalk, Gleser, Magliocco, and D'Zmure (1961).

The anxiety scale of Gleser, et al. (1961) - to be used in this study - is discussed in more detail together with supporting validity and reliability data in another

section of this report.

An application of the different methods, tools, and systems of content-analysis by Strupp, Jaffe, Mahl, and Gottschalk and his associates is found in the comparative analysis of two psychiatric interviews (Gottschalk, 1961).

In summary, a review of the various areas of research of speech and its related fields would seem to indicate that an analysis of continuous speech employing an electro-acoustic technique appears to be feasible after the ground work has been laid which encompasses electroacoustic studies of monosyllabic speech responses, filtering of speech, as well as a list of other studies which depended on listener-judgments for evaluation.

CHAPTER III

DEFINITIONS AND HYPOTHESES

DEFINITIONS AND HYPOTHESES

I Definitions

The following definitions have been adopted for use in this study. While some of them are generally accepted, others are defined operationally for the particular use in this investigation. In the majority of cases, an abbreviated form of the special terminology has been constructed which is used throughout the remainder of the study.

Sound Pressure: the amount of force acting over a unit area of surface measured in dyne per square centimeter (dyne/cm^2).

Intensity: the correlate of physical energy measured in Watt per $meter^2$ (Watt/m^2).

Frequency: the number of cycles or complete alternations per unit of time of a wave; measured in cycles per second or Hertz (Hz).

Broad Frequency Band: (BFB), frequencies from 80 - 6300 Hz.

First Formant Frequency Band: (FFFB), frequencies from 200 - 800 Hz.

Narrow Frequency Band: (NFB), frequencies from 80 - 250 Hz.

Decibel: (dB), a unit used to compare two voltages or currents equal to 20 times the common logarithm of the ratio of the voltage or currents measured across equal resistances.

Mean Speech Pressure: (MSP), the power ratio between mean voltages of appropriate blocks of speech samples multiplied by 20 times the common logarithm of the ratio of the voltages measured across equal resistances, expressed in decibels.

Spontaneous Speech Sample: (SSS), a sample of unrehearsed speech elicited according to a standard instruction.

Transient Emotional Sequence: (TES), a temporary physiologic condition experienced affectively rather than intellectually; identified and expressed in terms of scores derived from the GLESER anxiety scale.

Speech Sample: (SS), a clause or sentence or a combination of both of 10 to 15 seconds in duration.

Neutral Speech Sample: (NSS), a clause or sentence or a combination of both, judged to be neutral or lacking emotional content. NSS's are prepared samples of speech to be read aloud by the subject.

Emotional Speech Sample: (ESS), a clause or sentence or a combination of both, judged to contain emotional content. The ESS are prepared samples to be read by each subject.

Anxiety-free Speech Sample: (AfSS), a clause or sentence or a combination of both obtained from the sample of spontaneous speech. The AfSS was judged to be free of anxiety in as far as it did not receive any anxiety score.

Anxiety-Speech-Sample: (ASS), a clause or sentence or a combination of both obtained from the spontaneous speech sample. The ASS carries emotional content as measured by the GLESER anxiety scale.

Treatment Order: sequence of passage presentations having the following abbreviations

NES - NSS, ESS, SSS
NSE - NSS, SSS, ESS

ENS - ESS, NSS, SSS
ESN - ESS, SSS, NSS

SNE - SSS, NSS, ESS
SEN - SSS, ESS, NSS

II Hypotheses

The hypotheses stated below are derived from research reported in previous chapters which suggests that changes in the level of emotional arousal during speech production can be detected through changes in vocal intensity in certain frequency bands. The Broad Frequency Band (BFB) and the Narrow Frequency Band (NFB) have been investigated before to some extent using monosyllabic speech responses only. No similar investigation has been apparently performed on the bandwidth ranging from 200 - 800 Hz which comprises the first formant frequencies of all vowel phonemes. This vowel resonance region contains the greatest speech power.

Primary Hypotheses

For each bandwidth investigated, one major hypothesis and four sub-hypotheses are tested.

Hypothesis I: Frequency Bandwidth 80 - 250 Hz. Changes in the level of arousal during continuous speech production will result in concomitant changes in vocal intensity within this frequency range, while parity of levels of emotional arousal remains unaffected.

(a) There is a difference in MSP between NSSs and ESSs.

(b) There is a difference in MSP between AfSSs and ASSs.

(c) There is NO difference in MSP between ESSs and ASSs.

(d) There is NO difference in MSP between NSSs and AfSSs.

Hypothesis II: Frequency Bandwidth 200 - 800 Hz. Changes in the level of arousal during continuous speech production will result in concomitant changes in vocal intensity within this frequency range, while parity of levels of emotional arousal remains unaffected.

(a) There is a difference in MSP between NSSs and ESSs.

(b) There is a difference in MSP between AfSSs and ASSs.

(c) There is NO difference in MSP between ESSs and ASSs.

(d) There is NO difference in MSP between NSSs and AfSSs.

Hypothesis III: Frequency Bandwidth 80 - 6300 Hz. Changes in the level of arousal during continuous speech production will result in concomitant changes in vocal intensity within this frequency range, while parity of levels of emotional arousal remains unaffected.

(a) There is a difference in MSP between NSSs and ESSs.

(b) There is a difference in MSP between AfSSs and ASSs.

(c) There is NO difference in MSP between ESSs and ASSs.

(d) There is NO difference in MSP between NSSs and AfSSs.

These hypotheses were tested in identical fashion for each of the two groups of subjects in this study.

Secondary Hypotheses

In addition to the major intent of this investigation, four other hypotheses were tested which are identical for each one of the three bandwidths employed. The Mean Speech Intensities of neutral (NSS), emotional (ESS), anxiety-loaded (ASS), and anxiety-free (AfSS) speech samples produced by the NON-CLINICAL group are compared to those obtained from identical speech samples of the CLINICAL group. These secondary hypotheses, because they are not couched in a detailed rationale based on theory or research, are not stated in formal hypothesis form.

CHAPTER IV

METHOD

METHOD

I The Sample

The total sample consisted of 36 male subjects between twenty and forty years of age. This comprised two groups of subjects, one being drawn from the undergraduate population of the University of Alberta, the other, from the male workers of the Sheltered Workshop in Edmonton. Hereafter, the former will be referred to as NON-CLINICAL, the latter as CLINICAL.

All subjects in the NON-CLINICAL group were either in their first or second year of study. Their help for this study was solicited through the seminar leaders in the undergraduate courses ED PSY 269 and ED PSY 271.

The members of the CLINICAL group were discharged former mental patients of the Alberta Hospital, Edmonton. At the time of the study, all of them were living in foster homes in the City of Edmonton. These subjects were also chosen on the basis of their willingness to participate. The names of these subjects were submitted to the Alberta Hospital, Edmonton, in order to obtain a rating of the effects of their medication on perception, physiological reaction, and speech. Ratings of effects of medication on perception, physiological reaction, and speech are reproduced in Table I.

The rationale for the inclusion of the CLINICAL group was to investigate the potential applicability of this type of analysis of therapeutic interviews of people who have received intensive treatment in the past and were still receiving some degree of treatment at the time of the study. It was also assumed that subjects in the CLINICAL group manifest a difference in their vocal reactions compared to the NON-CLINICAL which is worth investigating.

TABLE I

EFFECTS OF MEDICATION

Degree of Effect	Perception	Physiological Reaction	Speech
None	5	2	3
Light	13	14	13
Medium		2	1
Severe			1
Total	18	18	18

In Table I, the CLINICAL subjects are grouped on a four-point scale according to the expected effects of their medication on perception, physiological reaction, and speech. As can be observed from this table, only one subject has been rated as suffering severe effects of medication during speech production. The great majority received 'none' or 'light' ratings on the three effect areas.

During the planning stages of this study, attempts were made to obtain two groups of subjects with approximately equal age distributions. As Table II shows, the age distributions of the two samples range between twenty to forty years of age, with a mean age of 27.06 years and a standard deviation of 5.94 for the NON-CLINICAL group. The mean age and standard deviation for the CLINICAL group are 30.28 years and 5.84 respectively. The NON-CLINICAL group comprises fewer students over the age of thirty while the CLINICAL group age distribution is equal between those over and under thirty years of age. The ages for both groups of subjects are reproduced in Table II below.

TABLE II

AGES OF SUBJECTS IN SAMPLES

Age	NON-CLINICAL frequency	total	CLINICAL frequency	total
20 - 21	4		1	
22 - 23	3		2	
24 - 25	2	12	1	9
26 - 27	1		4	
28 - 29	2		1	
30 - 31	1		1	
32 - 33	2		2	
34 - 35	1	6	2	9
36 - 37	1		1	
38 - 39	1		1	
40			2	
Total		18		18

II Testing Instruments

The instruments used in this study consisted of selected reading passages (NSS) judged by a body of students to contain no obvious emotional content, and of emotional passages (ESS) selected by the writer from diverse sources. Another instrument employed was the Gleser, et al., (1961) anxiety scale which is applicable to the analysis of transient emotional states in spoken or written material.

Neutral Speech Samples (NSS)

Thirty reading passages of approximately ten to fifteen seconds duration of oral reading time were selected by the writer from prose writings of various subject matter orientations including Geography, Agriculture, Biology, Theology, and Philosophy. These passages were rated by 43 males on a three-point rating scale according to the amount of emotional reaction experienced while reading them. Of the eight passages which consistently received the lowest rating, six were selected. These NSS were then printed in large letters on 5 x 8 inch index cards.

Emotional Speech Samples (ESS)

The six ESS used were selected out of a total of twenty-five reading passages drawn for their open sexual connotation from such prose writings as FANNY HILL, LADY CHATTERLEY'S LOVER, and PSYCHO-SEXUAL PROBLEMS by Szednik

(1964). The chosen six ESS received the highest ratings from the same group of males mentioned above. Most passages were modified in such a way that the subject read them as either narrator or active agent. The ESS like the NSS were printed in large letters on 5 x 8 inch index cards.

With regard to the use of reading passages of sexual content, it was assumed that sexual stimuli constitute an arousing situation to the individual. Byrne (1961) found that the experimental arousal of affiliation needs produced anxiety. In a more recent article, Byrne and Sheffield (1965) make the assumption that sexual stimuli would have an even greater arousal value than affiliation needs in our culture. The sexual stimuli employed in their study consisted of selected paragraphs from such novels as THE NAKED AND THE DEAD, ULYSSES, PEYTON PLACE, THE REVOLT OF MAMIE STOVER, and others. Another team of researchers, Aronson and Mills (1959), used twelve obscene words as well as vivid descriptions of sexual activity from contemporary novels in an 'embarrassment test' which was one of the instruments used in their study of severity of initiation rites and the subsequent liking for the group.

The selection of prepared passages of such different emotional contents as in the NSS and ESS has an advantage over spontaneous passages constructed by the subject himself in that word selection and combination is deter-

minable by the experimenter and is the same for all subjects. It is thought that this allows for a more exact comparison of degrees of physiological arousal between passages as well as between subjects.

Anxiety Speech Samples (ASS) and Anxiety-free Speech Samples (AfSS)

Two ASSs and two AfSSs of about ten to fifteen seconds speaking time were selected from the Spontaneous Speech Sample (SSS) of each subject after application of the GLESER Anxiety Scale. The ASSs constitute those passages which received anxiety scores, while the AfSSs are passages which are anxiety-free according to the theory of this scale. Samples of ASSs and AfSSs are reproduced together with all of the NSS and ESS used in this study in Appendix A.

The authors, Gleser, et al., (1961) describe the theory behind the scale in the following way:

> The type of anxiety which we are attempting to measure is what might be termed "free" anxiety in contrast to "bound" anxiety which manifests itself in conversion and hypochondriacal symptoms, in compulsions, in doing and undoing, in withdrawal from human relationships, and so forth. Our index of free anxiety might be expected to correlate with the clinical assessment of manifest anxiety, but our construct differs somewhat from that of manifest anxiety, in that it includes only psychological manifestations and not the autonomic and nonverbal total behavioral manifestations of anxiety (p. 595).

The assumptions underlying the scale are these: (1) statements of feelings of a certain type are considered reflecting the same amount of anxiety whether they are stated in the past, present or future tense; (2) statements about threatening incidents will be related with frequent self-references when most anxiety causing, when anxiety is somewhat less it may be expressed more indirectly through externalization or displacement using animate objects. Anxiety of a still lesser degree may be reported using inanimate objects. A denial of anxiety indicates its actual existence. (3) the greater the anxiety of the speaker, the more numerous will be the references to the type of anxiety experienced.

The GLESER Anxiety Scale

(1) Death anxiety - references to death, dying threat of death, or anxiety about death experienced by or occurring to:

a) self	(3)
b) animate others	(2)
c) inanimate objects destroyed	(1)
d) denial of death anxiety	(1)

(2) Mutilation (castration) anxiety - references to injury, tissue, or physical damage, or anxiety about injury or threat of such experienced by or occuring to:

a) self	(3)
b) animate others	(2)
c) inanimate objects	(1)
d) denial	(1)

(3) Separation anxiety - references to desertion, abandonment, ostracism, loss of support, falling, loss of love or love object, or threat of such experienced by or occurring to:

a) self	(3)
b) animate others	(2)
c) inanimate objects	(1)
d) denial	(1)

(4) Guilt anxiety - references to adverse criticism, abuse, condemnation, moral disapproval, guilt, or threat of such experienced by:

a) self	(3)
b) animate others	(2)
c) denial	(1)

(5) Shame anxiety - references to ridicule, inadequacy, shame, embarrassment, humiliation, overexposure of deficiencies or private details, or threat of such experienced by:

a) self	(3)
b) animate others	(2)
c) denial	(1)

(6) Diffuse or nonspecific anxiety - references by word or phrase to anxiety and or fear without distinguishing type or source of anxiety.

a) self	(3)*
b) animate others	(2)*
c) denial	(1)*

*The numbers in parentheses are the scoring weights for the different levels of anxiety.

Notes About the Scale

(1) The grammatical clause is the unit for scoring; exluded are expletives or elliptical expressions.

(2) Clauses which indicate that the speaker is the agent producing injury or expressing criticism, etc. directed

toward others are not scored.

(3) The weight of a score is increased by ONE if the statement of anxiety or fear is verbally modified to indicate that the condition is extreme or marked.

(4) Any grammatical form of the word is scored regardless of its gremmatical property.

(5) The Spontaneous Speech Passages (SSS) which were analyzed for anxiety content using the above anxiety scale were elicited using the following instructions:

> This is a study of speaking and conversational habits, I would like you to start telling me as freely as you can about any interesting or dramatic life experience you have had. Once you have started, I shall be here listening to you, but I would prefer not to reply to any questions you may feel like asking me until the FIVE minutes are up. Do you have any questions you would like to ask me now before we start?... Well, then, you may start upon my hand signal.

The above instruction was printed in large letters on a 5 x 8 inch index card and presented to the subject in the same form as the NSS and ESS.

(6) The magnitude of an emotion as measured by this scale can be expressed per hundred words spoken according to this formula:

$$\frac{100\ (f_1w_1 + f_2w_2 + \ldots f_nw_n + 0.5)}{N}^{\frac{1}{2}}$$

where f_n is the frequency per unit of time of any relevant type of thematic reference; w_n is the weight applied to

such verbal statements; and N is the number of words per unit time.

The results observed between the two groups of subjects, Table III, on the GLESER anxiety scale reflect the findings of Gleser's normative sample of 94 subjects and his sample of 24 psychiatric patients, respectively. Significant differences in mean anxiety scores at the 0.05 level, two-tailed, were found after application of a t-test.

TABLE III

DISTRIBUTION OF TRANSFORMED VERBAL ANXIETY SCORES IN THE TWO SAMPLES

Square Root of Score	Non-Clinical f	Clinical f
5.50 - 5.99		1
5.00 - 5.49		2
4.50 - 4.99	2	3
4.00 - 4.49	2	8
3.50 - 3.99	2	2
3.00 - 3.49	3	1
2.50 - 2.99	4	
2.00 - 2.49	3	1
1.50 - 1.99		
1.00 - 1.49		
0.50 - 0.99		
0.00 - 0 49	2	
N	18	18
$\bar{x}$	2.89	4.27
S.D.	1.27	0.78

Validity of the GLESER Anxiety Scale

The following data on validity have been reported by the authors of the scale, Gleser, et al., in their 1961 publication: (a) a comparison of distributions of anxiety scores between 24 psychiatric patients and 94 normals resulted in a difference in means which was significant beyond the 0.001 level. (b) independent ratings of 24 subjects using a clinical scale of anxiety administered by two resident psychiatrists - immediately preceding the giving of the five-minute speech sample by the subjects - and scores obtained from the GLESER Anxiety Scale resulted in a Pearson product-moment correlation of 0.66 which is significant beyond the 0.001 level. (c) a comparison between the MMPI Pt Scale, corrected for K, and the GLESER scale resulted in a correlation of 0.51. (d) in still another study involving fourteen dermatological patients, a comparison between the WELSH A-Scale of the MMPI and the GLESER scale produced a correlation of 0.68, which is significant beyond the 0.01 level. Gottschalk, Kaplan, Gleser, and Winget (1962) reported that the total anxiety score for three of five women was lowest during the 'low hormone' phase and reached statistical significance. These subjects had been requested to deliver 5-minute verbal samples three to seven days per week covering time before and after their menses.

Scoring Reliability of the GLESER Anxiety Scale

Gleser, et al., (1961) reported a correlation coefficient of 0.86 after scoring of twenty protocols of the normative sample by two independent scorers. Rank-order correlations for the six sub-categories were: I (death) 0.44; II (mutilation) 0.94; III (separation) 0.68; IV (guilt) 0.83; V (shame) 0.83; and VI (diffuse) 0.75. The scoring of 17 protocols - all from the same subject - by three coders resulted in a correlation of 0.83 for the total score coded by any scorer. The average reliability resulting out of the same analysis of variance for any of the subscales using a single coder was 0.68. The recoding of the above protocols of the normative study after one year and an analysis of variance application to these scores, with neither significant interaction nor significant main effects for coder over time, resulted in a reliability estimate of 0.76. The authors state that this is a 'conservative' measure which might be expected when using randomly selected trained technicians and scoring at different times. The overall reliability of two independent scorings would result in a minimum reliability of 0.86 for the total anxiety score.

III Procedure of Recording

Tape recording took place in a sound-insulated room in the Department of Speech Pathology and Audiology. The subject was seated against a table facing the only window in this room. This arrangement allowed subject and experimenter to view each other during the recording session.

In order to keep the subject within a specified distance (15 inches) from the microphone, which was placed on the table before him, a cloth-covered, dowel-mounted wire loop (see Appendix B) attached to an adjustable microphone stand was placed over the subject's head. The wire loop was sufficiently large and mouldable to fit over the subject's head without actually touching it.

Before the actual recording, the subject was told that he could communicate with the experimenter via an intercom system. Next, he was instructed to read a passage printed in large letters on a 5 x 8 inch card in order to set the volume on the tape recorder. All reading passages were presented on cards of the above dimensions. These cards were placed at eye-level into a window-mounted paper envelope which was within reading distance from the subject. Each subject was questioned whether or not he could see and read the print at ease. One subject belonging to the CLINICAL group had to be rejected because of his inability to read the passages in the accepted form. He was partially

blind on one eye, and in need of glasses for the other.

After the above preparations, each subject was instructed according to treatment order to either read the NSS, the ESS, or to produce up to five minutes of Spontaneous Speech (SSS). The presentation and recording of NSS, ESS, and SSS were alternated with three subjects following one particular treatment order within each of the two groups of subjects. Each subject was given time to read each of the passages once silently before reading it aloud. Prior to the recording proper, each subject spoke a prepared identification number on tape also indicating treatment order.

IV Processing of Recorded Voice Samples

The original recording was made on an AMPEX AG-500 tape recorder at 7½ ips using an ELECTRO-VOICE microphone, Model 666. After the SSSs were transcribed verbatim using a transcribing tape recorder, the writer scored these protocols and selected several ASSs and AfSSs from each protocol. These selected passages, together with the complete protocols, were submitted to a co-rater, Mr. W. Green, a graduate student in the Department of Educational Psychology, for an independent re-scoring. Only those passages on which agreement was obtained were selected for further analysis. A correlation coefficient of 0.83 was obtained between the independent scorings of writer and

co-rater.

Two subjects belonging to the NON-CLINICAL group produced SSSs which contained no scorable material. Two randomly selected passages were chosen as substitutes from the protocols of each of the subjects to enter the statistical analysis.

Specifications of the different types of equipment used together with a diagrammatical sketch of the circuitry of the equipment are found in another part of this study (See Appendix C).

A set of Master tapes were produced containing the NSSs, ESSs, ASSs, and AfSSs of all subjects. The tape-recorders used in the production of these Master tapes were the AMPEX AG-500, on which the original tapes were made, and an AMPEX AV-770.

In order to quantify the vocal intensities during speech production for each one of the three bandwidths used, the signal from the pre-amplifier of the AMPEX AV-770 tape recorder was passed through a band-pass filter (SKL Variable Electronic Filter, Model 308 A) into a RMS Voltmeter (HEWLETT-PACKARD Model 3400 A) and from there into a Data Acquisition System (HEWLETT-PACKARD Model 2010 J) which produced a record of the relative amplitude of acoustic energy on a digital magnetic tape. The resolution time of the Data System was set at 0.1 sec.

which means that the incoming signal is sampled at a rate of six times per second. NON-CLINICAL subjects were recorded in the 0.3 volts sensitivity range, while CLINICAL subjects required a 0.03 volts setting. The data produced in this way can be entered into a computer for printout and further analysis.

In order to achieve satisfactory spacing of passages during the processing described above, all passages were spaced 40 seconds apart on the Master tapes. Since the actual length of each passage was less than half that time, but signal sampling of the Data Acquisition System was set at 214 samples, amounting to a slightly less than 40 seconds, noise of the tape recorder accounted for one third of the readings on the magnetic tape. After an intensive study of the recorded readings on paper print-outs, cut-off points were chosen in order to drop the superfluous tape recorder noise. The remaining readings constitute the actual equivalents of vocal intensities during speech production. In this type of processing, the relative amplitudes of intensities are expressed in units of millivolts. The MSP was computed using the mean readings in millivolts between blocks of speech passages. These values were used in the statistical analysis.

CHAPTER V

RESULTS

RESULTS

I Statistical Analysis

In order to test the hypotheses stated in Chapter III, an analysis was required which would compare the differences in Mean Speech Pressure (MSP) between the appropriate blocks of speech samples separately over each frequency bandwidth for each one of the six treatments employed in this study. For this purpose a two-factor analysis of variance (Winer, 1962) with repeated measures on one factor (frequency bandwidth) was chosen. This design permits individual comparisons of variability in scores for each treatment over the repeated factor.

To test the effects of treatment order as well as those of bandwidth an F-ratio is computed which signifies at what level of significance treatment or bandwidth or a combination of both affect the differences of MSP between the appropriate blocks of speech samples. For the purpose of this study, the level of significance was set at the 0.05 level. In order to test for differences between all possible pairs of treatment group means as well as between bandwidth means the Newman-Keuls procedure was adopted which permits the testing for statistical significance between means 'r' steps apart on an ordered scale.

II Findings

The tables below present the results of the two-factor experiment with repeated measures. Each one of the analysis of variance tables is succeeded by a second table providing the individual treatment means as well as those for each frequency bandwidth analyzed. Where applicable the results of tests of differences between ordered means are given. Findings are presented separately for the two groups studied, with the NON-CLINICAL group preceding the CLINICAL one. Table XII provides a summary of the analyses of variance and ordered means for the NON-CLINICAL group. Table XIII serves the same purpose for the CLINICAL group. A discussion of the relationship between findings and each of the three major hypotheses follows each table.

Findings, NON-CLINICAL Group

In the next set of tables the results of the summaries of analyses of variance are reproduced together with summary tables of cell means for each analysis of variance.

The results of Table IV (page 58) indicate that the order of treatment has a significant effect on the difference in MSP between NSS and ESS. Since the second main effect, bandwidth, is not significant, but the AxB interaction is, tests on simple main effects are called for.

TABLE IV

NON-CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCE IN MSP BETWEEN NSS AND ESS SPEECH SAMPLES

Source of Variation	SS	df	MS	F	P
Between	8.431	17			
Treatment (A)	5.153	5	1.031	3.773	0.027
Sub. w.g.	3.278	12	0.273		
Within	8.083	36			
Bandwidth (B)	0.799	2	0.400	2.562	0.098
A x B	3.540	10	0.354	2.269	0.048
B x Sub. w.g.	3.744	24	0.156		

A test on the simple main effects of factor B (bandwidth) at each treatment level showed that for bandwidth 1 and treatment order ESN, the MSP is significantly lower than for treatments NES and ENS. No differences were found among the different treatments in bandwidth 2. In bandwidth 3, the differences between treatment orders NES and ESN reached statistical significance. The latter treatment order has the lowest MSP. Mean Speech Pressure values for each bandwidth and treatment order are given in Table V on page 59.

TABLE V

NON-CLINICAL
DIFFERENCES IN MSP BETWEEN NSS AND ESS FOR
EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth 1 80-250 Hz	2 200-800 Hz	3 80-6300 Hz	Grand Means of Treatment
N E S	-1.919	-1.319	-1.752	-1.663
N S E	-1.006	-0.697	-0.645	-0.783
E N S	-1.696	-1.242	-1.282	-1.407
E S N	-0.629	-1.576	-0.419	-0.875
S N E	-1.186	-1.174	-0.773	-1.044
S E N	-1.099	-1.658	-1.186	-1.314
Means of Frequency Bandwidths	-1.422	-1.278	-1.009	

A summary of the analysis of variance in Table VI (page 60) on the differences in MSP between anxiety free and anxiety loaded speech samples shows that bandwidth has a significant effect on the MSP between the two groups of samples compared. Neither treatment order nor interaction between the former and bandwidth were significant. The results of a Newman-Keuls comparison between ordered means of factor B (bandwidth) indicated that bandwidth 3 is significantly different from bandwidth 2. No difference was observed between bandwidths 1 and 2.

TABLE VI
NON-CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCES
IN MSP BETWEEN AfSS AND ASS

Source of Variation	SS	df	MS	F	P
Between	16.453	17			
Treatment (A)	5.078	5	1.016	1.072	0.422
Sub. w.g.	11.375	12	0.948		
Within	10.798	36			
Bandwidth (B)	3.494	2	1.747	6.549	0.005
A x B	0.902	10	0.090	0.338	0.961
B x Sub. w.g.	6.402	24	0.267		

TABLE VII
NON-CLINICAL
DIFFERENCES IN MSP BETWEEN AfSS AND ASS
FOR EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth 1 80-250 Hz	2 200-800 Hz	3 80-6300 Hz	Grand Means of Treatment
N E S	-0.817	-1.355	-0.621	-0.931
N S E	-0.569	-0.633	-0.451	-0 551
E N S	-1.325	-1.523	-0.773	-1.207
E S N	-1.570	-1.518	-1.108	-1.399
S N E	-0.949	-1.290	-0.249	-0.829
S E N	-1.403	-1.649	-1.075	-1.376
Means of Frequency Bandwidths	-1.105	-1.328	-0.713	

As the below summary, Table VIII, of analysis of variance shows none of the main effects reached the required level of significance. The probability level for the A x B interaction is equally low and does not warrant testing for simple main effects of either A or B.

TABLE VIII

NON-CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCE
IN MSP BETWEEN ESS AND ASS SPEECH SAMPLES

Source of Variation	SS	df	MS	F	P
Between	453.201	17			
Treatment (A)	68.242	5	13.648	0.425	0.822
Sub. w.g.	384.959	12	32.080		
Within	94.618	36			
Bandwidth (B)	8.336	2	4.168	1.555	0.231
A x B	21.946	10	2.195	0.819	0.614
B x Sub. w.g.	64.335	24	2.681		

TABLE IX

NON-CLINICAL
DIFFERENCES IN MSP BETWEEN ESS AND ASS
FOR EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth 1 80-250 Hz	2 200-800 Hz	3 80-6300 Hz	Grand Means of Treatment
N E S	-0.176	+0.405	-0.387	-0.052
N S E	-1.683	-1.721	-1.997	-1.800
E N S	-0.165	-3.306	-3.504	-2.325
E S N	-3.542	-3.165	-3.264	-3.324
S N E	-0.360	+0.016	-0.730	-0.358
S E N	-0.706	-2.709	-2.404	-1.940
Means of Frequency Bandwidths	-1.105	-1.747	-2.048	

The results of the following analysis of variance, Table X, indicate that neither main effects nor interaction affects significantly effected the difference in MSP between Neutral Speech Samples and Anxiety-free Speech Samples.

TABLE X

NON-CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCE
IN MSP BETWEEN NSS AND AfSS SPEECH SAMPLES

Source of Variation	SS	df	MS	F	P
Between	618.984	17			
Treatment (A)	82.908	5	16.582	0.371	0.858
Sub. w.g.	536.076	12	44.673		
Within	68.424	36			
Bandwidth (B)	10.737	2	5.368	3.008	0.068
A x B	14.845	10	1.485	0.832	0.603
B x Sub. w.g.	42.841	24	1.785		

TABLE XI

NON-CLINICAL
DIFFERENCES IN MSP BETWEEN NSS AND AfSS FOR
EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth 1 80-250 Hz	2 200-800 Hz	3 80-6300 Hz	Grand Means of Treatment
N E S	-0.907	-0.951	-2.255	-1.371
N S E	-1.350	-1.453	-1.381	-1.395
E N S	+0.921	-0.710	-1.454	-0.414
E S N	-2.832	-3.805	-3.683	-3.440
S N E	+0.014	-0.035	-2.334	-0.785
S E N	+0.591	+0.319	+0.996	+0.635
Means of Frequency Bandwidths	-0.594	-1.106	-1.685	

Findings, CLINICAL Group

The findings for this group are reproduced in the following set of tables which provide summaries of analyses of variance as well as cell means for each analysis.

TABLE XII

CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCE IN MSP BETWEEN NSS AND ESS SPEECH SAMPLES

Source of Variation	SS	df	MS	F	P
Between	9.732	17			
Treatment (A)	6.666	5	1.333	5.218	0.008
Sub. w.g.	3.066	12	0.255		
Within	21.343	36			
Bandwidth (B)	10.463	2	5.231	20.881	0.000
A x B	4.868	10	0.487	1.943	0.088
B x Sub. w.g.	6.013	24	0.251		

Since the two main effects were both significant, a Newman-Keuls procedure for making comparisons between all possible pairs of ordered means was performed for the treatment order effect as well as for the second main effect, bandwidth. With respect to treatment order it was found that treatments NSE, ESN and SNE differed significantly from treatment SEN. A test on main effect B showed that bandwidths 1 and 2 are significantly different from bandwidth 3 which had the lowest difference

in MSP between Neutral and Emotional Speech Samples.

TABLE XIII

CLINICAL
DIFFERENCES IN MSP BETWEEN NSS AND ESS FOR
EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth 1 80-250 Hz	2 200-800 Hz	3 80-6300 Hz	Grand Means of Treatment
N E S	-1.148	-0.839	-0.214	-0.733
N S E	-1.915	-1.887	-0.316	-1.372
E N S	-1.436	-1.014	-0.421	-0.957
E S N	-1.572	-2.052	-0.336	-1.320
S N E	-1.537	-1.619	-0.437	-1.197
S E N	-0.333	-0.298	-0.502	-0.377
Means of Frequency Bandwidths	-1.323	-1.285	-0.371	

The summary of the analysis of variance in Table XIV on page 66 shows that only main effect B, bandwidth, has a significant effect on the MSP between the two groups of speech samples compared. The application of a Newman-Keuls test on ordered means of main effect B produced significant results for bandwidths 2 and 3. Table XV shows that bandwidth 1 has the lowest MSP for the speech samples analyzed.

TABLE XIV

CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCES IN MSP BETWEEN AfSS AND ASS SPEECH SAMPLES

Source of Variation	SS	df	MS	F	P
Between	24.372	17			
Treatment (A)	13.529	5	2.706	2.994	0.065
Sub. w.g.	10.843	12	0.904		
Within	18.831	36			
Bandwidth (B)	6.366	2	3.183	11.091	0.000
A x B	5.577	10	0.558	1.944	0.088
B x Sub. w.g.	6.887	24	0.287		

TABLE XV

CLINICAL
DIFFERENCES IN MSP BETWEEN AfSS AND ASS FOR EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth			Grand Means of Treatment
	1 80-250 Hz	2 200-800 Hz	3 80-6300 Hz	
N E S	-1.179	-2.439	-2.206	-1.941
N S E	-2.089	-2.804	-1.628	-2.173
E N S	-0.546	-0.709	-1.616	-0.957
E S N	-0.505	-1.133	-0.999	-0.879
S N E	-1.165	-1.611	-2.248	-1.675
S E N	-0.336	-1.576	-1.404	-1.105
Means of Frequency Bandwidths	-0.970	-1.712	-1.683	

TABLE XVI

CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCES
IN MSP BETWEEN ESS AND ASS SPEECH SAMPLES

Source of Variation	SS	df	MS	F	P
Between	285.413	17			
Treatment (A)	139.630	5	27.926	2.299	0.110
Sub. w.g.	145.783	12	12.149		
Within	103.299	36			
Bandwidth (B)	25.139	2	12.569	5.215	0.013
A x B	20.312	10	2.031	0.843	0.594
B x Sub. w.g.	57.848	24	2.410		

TABLE XVII

CLINICAL
DIFFERENCES IN MSP BETWEEN ESS AND ASS
FOR EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth			Grand Means of Treatment
	1	2	3	
	80-250 Hz	200-800 Hz	80-6300 Hz	
N E S	-1.405	+0.750	+0.205	-0.150
N S E	-2.788	-1.096	-2.369	-2.084
E N S	-0.884	+1.925	+3.003	+1.348
E S N	-2.357	-2.328	-2.070	-2.252
S N E	-4.995	-3.294	-2.659	-3.649
S E N	-1.067	-0.982	-0.722	-0.924
Means of Frequency Bandwidths	-2.249	-0.838	-0.769	

A comparison on the difference between ordered means, Table XVI, page 67, for the significant main effect B, bandwidth, produced results which indicate that bandwidth 1 is significantly different from bandwidth 2 and 3. The difference between the latter bandwidths is not significant.

TABLE XVIII

CLINICAL
SUMMARY OF ANALYSIS OF VARIANCE OF DIFFERENCES
IN MSP BETWEEN NSS AND AfSS SPEECH SAMPLES

Source of Variation	SS	df	MS	F	P
Between	231.221	17			
Treatment (A)	56.924	5	11.385	0.784	0.580
Sub. w.g.	174.296	12	14.525		
Within	144.285	36			
Bandwidth (B)	30.837	2	15.418	4.875	0.016
A x B	37.535	10	3.754	1.187	0.346
B x Sub. w.g.	75.913	24	3.163		

Since main effect B, bandwidth, is significant in the above summary of analysis of variance, a Newman-Keuls comparison between all possible pairs of means was performed. The application of this test showed that bandwidth 3 is significantly different from the other two bandwidths, as can be observed from Table XIX.

TABLE XIX

CLINICAL
DIFFERENCES IN MSP BETWEEN NSS AND AfSS FOR
EACH TREATMENT ORDER AND BANDWIDTH

Treatment Order	Frequency Bandwidth 1 80-250 Hz	2 200-800 Hz	3 80-6300 Hz	Grand Means of Treatment
N E S	-2.014	-2.844	+1.464	-1.131
N S E	-1.753	-2.291	+1.631	-0.804
E N S	+1.439	+1.644	+1.364	+1.482
E S N	-2.277	-1.056	-0.144	-1.159
S N E	-1.189	-0.098	-0.255	-0.514
S E N	+0.495	+1.199	+1.051	+0.915
Means of Frequency Bandwidths	-0.883	-0.574	+0.851	

Findings of NON-CLINICAL Group (Summary)

The data presented in the preceding tables are summarized for each group of subjects in the summary table below. Table XX contains all relevant data concerning the NON-CLINICAL group of subjects. The data are given separately for each frequency bandwidth analyzed in order to facilitate the discussion of individual hypotheses.

TABLE XX

NON-CLINICAL
SUMMARY OF RESULTS OF ANALYSES OF
VARIANCE FOR ALL BANDWIDTHS

Speech Samples	Bandwidth Effect	Individual Treatments NES	NSE	ENS	ESN	SNE	SEN
B a n d w i d t h 80 - 250 Hz							
NSS - ESS	x	x		x	-		
AfSS - ASS	x	-	-	-	-	-	-
ESS - ASS	-	-	-	-	-	-	-
NSS - AfSS	-	-	-	-	-	-	-
B a n d w i d t h 200 - 800 Hz							
NSS - ESS	x	-	-	-	-	-	-
AfSS - ASS	x	-	-	-	-	-	-
ESS - ASS	-	-	-	-	-	-	-
NSS - AfSS	-	-	-	-	-	-	-
B a n d w i d t h 80 - 6300 Hz							
NSS - ESS	x	x			-		
AfSS - ASS	-	-	-	-	-	-	-
ESS - ASS	-	-	-	-	-	-	-
NSS - AfSS	-	-	-	-	-	-	-

x Difference significant at 0.05 level

\- Difference not significant

Hypothesis I

It was stated in this hypothesis that changes in level of arousal during continuous speech production will result in concomitant changes in MSP within the frequency range from 80 - 250 Hz. Analyses of its sub-hypotheses produced the following results.

(a) There is a difference in MSP between NSSs and ESSs.

This hypothesis was only partly confirmed since treatment order interacted with bandwidth. Two treatment orders, NES and ENS, were significantly different from a third order, ESN. The latter produced the lowest MSP.

(b) There is a difference in MSP between AfSSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(c) There is NO difference in MSP between ESSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(d) There is NO difference in MSP between NSSs and AfSSs. This hypothesis is confirmed at the 0.05 level.

Considering the results above, it seems tenable to accept Hypothesis I in principle for the NON-CLINICAL group of subjects. Treatment order affected the MSP between speech passages in that situation only in which order was most crucial. Treatment order ESN calls for two emotional

Hypothesis I

It was stated in this hypothesis that changes in level of arousal during continuous speech production will result in concomitant changes in MSP within the frequency range from 80 - 250 Hz. Analyses of its sub-hypotheses produced the following results.

(a) There is a difference in MSP between NSSs and ESSs.

This hypothesis was only partly confirmed since treatment order interacted with bandwidth. Two treatment orders, NES and ENS, were significantly different from a third order, ESN. The latter produced the lowest MSP.

(b) There is a difference in MSP between AfSSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(c) There is NO difference in MSP between ESSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(d) There is NO difference in MSP between NSSs and AfSSs. This hypothesis is confirmed at the 0.05 level.

Considering the results above, it seems tenable to accept Hypothesis I in principle for the NON-CLINICAL group of subjects. Treatment order affected the MSP between speech passages in that situation only in which order was most crucial. Treatment order ESN calls for two emotional

presentations prior to the production of neutral speech samples. A carry-over effect may be expected here.

Hypothesis II

This hypothesis stated that changes in the level of arousal during continuous speech production will result in concomitant changes in MSP within the frequency bandwidth from 200 - 800 Hz. The sub-hypotheses provided the results below.

(a) There is a difference in MSP between NSSs and ESSs.

(b) There is a difference in MSP between AfSSs and ASSs.

(c) There is NO difference in MSP between ESSs and ASSs.

(d) There is NO difference in MSP between NSSs and AfSSs.

The four preceding hypotheses are confirmed at the 0.05 level.

As can be observed from Table XX, this bandwidth produced significant differences in MSP between those blocks of speech samples in which variations could be expected. For this frequency range treatment order produced no interfering results for the NSS - ESS analysis. In view of these findings, hypothesis II appears to be tenable and is accepted.

Hypothesis III

This hypothesis is a replication of the previous two with the exception that it requires an analysis of the frequency bandwidth from 80 - 6300 Hz. Its sub-hypotheses produced the following results.

(a) There is a difference in MSP between NSSs and ESSs.

This hypothesis was only partly confirmed since treatment order interacted with bandwidth. Treatment order NES was significantly different from order ESN which produced the lower MSP.

(b) There is a difference in MSP between AfSSs and ASSs. This hypothesis is not confirmed.

(c) There is NO difference in MSP between ESSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(d) There is NO difference in MSP between NSSs and AfSSs. This hypothesis is confirmed at the 0.05 level.

Two of the above sub-hypotheses support the major ine in full and one supports it in part. Expressed in other words, hypotheses (c) and (d) state that speech samples with assumed equal degrees of emotionality do not differ in MSP. This assumption was confirmed. With respect to hypothesis (a), the data indicates that treatment order affects the MSP between the neutral and emotional speech

samples in a way similar to that observed in the analysis of bandwidth 1. The above results do not support the unanimous acceptance of this hypothesis.

Findings in CLINICAL Group (Summary)

Table XXI summarizes the relevant data of the preceding individual analyses of variance for this group of subjects. The data are presented separately for each frequency bandwidth.

Hypothesis I

It was hypothesized that changes in level of arousal during continuous speech production will result in concomitant changes in MSP within the frequency range from 80 - 250 Hz. The results of its four sub-hypotheses are given below.

(a) There is a difference in MSP between NSSs and ESSs.

This hypothesis is only partly confirmed since treatment order interacted with bandwidth. Three treatment orders, NSE, ESN, and SNE, were significantly different from order SEN which produced the lowest MSP.

(b) There is a difference in MSP between AfSSs and ASSs. This hypothesis is not confirmed.

(c) There is NO difference in MSP between ESSs and ASSs. This hypothesis is not confirmed.

(d) There is NO difference in MSP between NSSs and AfSSs. This hypothesis is confirmed at the 0.05

TABLE XXI

CLINICAL
SUMMARY OF RESULTS OF ANALYSES OF
VARIANCE FOR ALL BANDWIDTHS

Speech Samples	Bandwidth Effect	Individual Treatments NES	NSE	ENS	ESN	SNE	SEN
B a n d w i d t h 80 - 250 Hz							
NSS - ESS	x		x		x	x	-
AfSS - ASS	-	-	-	-	-	-	-
ESS - ASS	x	-	-	-	-	-	-
NSS - AfSS	-	-	-	-	-	-	-
B a n d w i d t h 200 - 800 Hz							
NSS - ESS	x	-	-	-	-	-	-
AfSS - ASS	x	-	-	-	-	-	-
ESS - ASS	-	-	-	-	-	-	-
NSS - AfSS	-	-	-	-	-	-	-
B a n d w i d t h 80 - 6300 Hz							
NSS - ESS	-	-	-	-	-	-	-
AfSS - ASS	x	-	-	-	-	-	-
ESS - ASS	-	-	-	-	-	-	-
NSS - AfSS	x	-	-	-	-	-	-

x Difference significant at 0.05 level.

\- Difference not significant.

level.

The results of the findings are at variance in certain cases with the hypothesized outcomes. The major hypothesis is supported in full by one sub-hypothesis only and in part by another. Two sub-hypothesis contradicted the expected direction. In view of these findings, hypothesis I appears not to be tenable and is therefore rejected.

Hypothesis II

This hypothesis is the same in content as the previous one but requires the analysis of the frequency bandwidth from 200 - 800 Hz. Tests on the data produced the following results in answer to the four sub-hypotheses below.

(a) There is a difference in MSP between NSSs and ESSs. This hypothesis is confirmed at the 0.05 level.

(b) There is a difference in MSP between AfSSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(c) There is NO difference in MSP between ESSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(d) There is NO difference in MSP between NSSs and AfSSs. This hypothesis is confirmed at the 0.05

level.

All four sub-hypotheses are in support of the major one. Therefore, hypothesis II appears to be tenable and is accepted.

Hypothesis III

This hypothesis with its respective four sub-hypotheses is identical in content with the previous ones but requires the analysis of the frequency bandwidth from 80 - 6300 Hz.

(a) There is a difference in MSP between NSSs and ESSs. This hypothesis is not confirmed.

(b) There is a difference in MSP between AfSSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(c) There is NO difference in MSP between ESSs and ASSs. This hypothesis is confirmed at the 0.05 level.

(d) There is NO difference in MSP between NSSs and AfSSs. This hypothesis is not confirmed.

While two sub-hypotheses support the major one, it is rejected by an equal number. In view of these findings, hypothesis III appears not to be tenable without restrictions and is therefore rejected.

Results of Secondary Hypotheses

These hypotheses required comparisons of Mean Speech Intensities of neutral (NSS), emotional (ESS), anxiety-loaded (ASS), and anxiety-free (AfSS) speech samples produced by the NON-CLINICAL group to those obtained from identical samples of the CLINICAL group. Results, for each of the three bandwidths, are presented in Table XXII below. As the results of these comparisons indicate, Mean Speech Intensities between the two groups of subjects are significantly different. The respective means values for the CLINICAL group are consistently lower than those of the NON-CLINICAL group.

III Conclusions

The foregoing analysis of the data obtained for this study would seem to suggest the following conclusions which are presented separately for each one of the groups studied in this investigation.

NON-CLINICAL

(a) Frequency Bandwidth 80 - 250 Hz. Changes in the level of arousal (AfSS - ASS) as well as parity between levels (ESS - ASS, and NSS - AfSS) is consistently reflected for comparisons between those blocks of speech samples for which treatment order is not crucial. In cases where

TABLE XXII

SUMMARY OF COMPARISONS OF MEAN SPEECH INTENSITIES OF SPEECH SAMPLES FOR NON-CLINICAL AND CLINICAL GROUP

Variable	NON-CLINICAL Mean	CLINICAL Mean	df	t Ratio	P two-tailed
Bandwidth 80 - 250 Hz					
NSS	0.11	0.00	34	5.293	0.000
ESS	0.08	0.00	34	8.184	0.000
ASS	0.09	0.00	34	4.228	0.000
AfSS	0.10	0.00	34	4.075	0.000
Bandwidth 200 - 800 Hz					
NSS	0.17	0.02	34	7.804	0.000
ESS	0.13	0.03	34	5.771	0.000
ASS	0.12	0.03	34	4.587	0.000
AfSS	0.17	0.03	34	5.804	0.000
Bandwidth 80 - 6300 Hz					
NSS	0.21	0.04	34	9.896	0.000
ESS	0.22	0.04	34	6.282	0.000
ASS	0.17	0.03	34	5.997	0.000
AfSS	0.21	0.05	34	7.525	0.000

both blocks of speech samples (NSS - ESS) are dependent on treatment order results are obtained which are at variance with the above statement.

(b) Frequency Bandwidth 200 - 800 Hz. This frequency range consistently reflects changes in level of arousal as well as parity between levels of arousal and is not affected by treatment order.

(c) Frequency Bandwidth 80 - 6300 Hz. This bandwidth does not adequately reflect changes in the level of arousal.

CLINICAL

(a) Frequency Bandwidth 80 - 250 Hz. This bandwidth does neither adequately nor consistently reflect changes in level of arousal or parity of arousal during continuous speech production.

(b) Frequency Bandwidth 200 - 800 Hz. This range consistently reflects changes in level of arousal as well as parity between levels of arousal during speech production. It is not affected by treatment order.

(c) Frequency Bandwidth 80 - 6300 Hz. This bandwidth does neither adequately nor consistently reflect changes in the level of arousal or parity of arousal during continuous speech production.

Summary of Conclusions

The three frequency bands employed in this investigation reflect changes in emotional arousal or parity of arousal during speech production with varying degrees of accuracy. The most accurate range, for both groups of subjects, was the 200 - 800 Hz range or bandwidth 2 followed by bandwidth 1 (80 - 250 Hz) and bandwidth 3 (80 - 6300 Hz). The Mean Speech Intensity of NON-CLINICAL subjects is significantly higher than that of CLINICAL subjects.

CHAPTER VI

DISCUSSION AND IMPLICATIONS

DISCUSSION AND IMPLICATIONS

I Discussion

In this study, an attempt was made to investigate the feasibility of using electroacoustic analyses of different frequency bandwidths in order to detect changes in transient, emotional states in continuous speech. It had been hypothesized, based on other research conducted in this area, that subjects change the intensity of their speech as they proceed from neutral to emotional material or vice versa. While studies of this nature have been reported using monosyllabic speech responses only, the present researcher felt that application of similar methods to continuous speech would greatly advance our understanding of expression of emotions or arousal through the speech channel. Of even greater interest to the researcher however, was his desire to find a somewhat more objective approach to the study of communication between client and therapist during therapeutic interviews.

A study was devised which required subjects to read factual or neutral passages as well as emotional ones into a high-fidelity tape-recorder. The use of reading passages had the advantage that the choice of words and the order of presentation were under direct control of the experimenter. In addition to these passages, subjects were

requested to deliver a five minute monologue on any dramatic event in their life. Factual and emotional passages were drawn from the transcribed monologue and sound-analyzed.

As the results and conclusions of this study already indicate, the different bandwidths investigated produced slightly differential results for the two groups of subjects employed. It can be stated, however, with some degree of certainty, that the bandwidth 80 - 6300 Hz, which comprised the essential range of speech for clear communication, is the least satisfactory bandwidth with regard to the electroacoustic analysis of arousal changes portrayed or reflected through speech behavior. While this is somewhat at variance with the findings of Rubenstein who reported statistically significant changes in monosyllabic speech responses between pre and post test of his experiment, note should also be taken of the fact that continuous speech is much more complicated as it involves an almost infinite number of word combinations which in themselve introduce changes into speech.

The 80 - 250 Hz frequency band, as has been reported earlier, has been investigated with considerable success, but again only monosyllabic speech responses were employed. The present study would indicate that compared to the broader band, the narrower range contains more information about the emotional state of the individual.

With respect to the NON-CLINICAL group, this bandwidth produced clear results across three sub-hypothesis, two of which stated that parity of levels of arousal results in insignificant differences in Mean Speech Pressure between blocks of passages. The third confirmed hypothesis stated that there is a difference in MSP between AfSSs and ASSs. Somewhat controversial results were obtained in the comparison between NSSs and ESSs. While treatment apparently did not affect arousal changes in certain sequences of treatment, other treatment orders did. Actually, only one treatment order resulted in insignificant MSP between passages. It seems logical to reason that a carry-over effect was present in this study. Supporting evidence for this assumption can be found in the study by Alpert and his associates who reported that differences between neutral and emotional stimuli were least significant after the presentation of the emotional stimulus. The carry-over effect wore off the further the trials were removed from the emotional stimulus presentation.

In the present study, the ESN treatment order produced the lowest MSP between NSSs and ESSs. From this treatment order it is obvious that two arousing situations preceded the final presentation of factual speech samples (NSS). On the average, the NSS succeeded the two earlier presenta-

tions by approximately one to one and a half minutes. It could be reasoned that the assumed carry-over effect would wear off as the presentation of NSSs came to conclusion. An investigation of the MSP for the latter passages showed a considerable increase in intensity. Therefore, if only the last two NSS passages had been employed in the analyses, the final results may indicate considerable changes in MSP between factual and emotional speech passages.

Slightly more controversial were the results in this band for the CLINICAL group of subjects, where it was the SEN treatment order which received the lowest MSP reading between passages. Here again, two emotional treatments precede the neutral presentation (NSS) producing apparently a high carry-over effect. It could be taken from the personal reports of subjects after the recording or the considerably higher anxiety scores of CLINICAL subjects that the five minute monologue was the most harrassing part of the experiment which affected the outcome of this particular treatment order.

The only frequency band which remained consistently unaffected by treatment order, not only for the NON-CLINICAL but also for the CLINICAL group of subjects was the range from 200 - 800 Hz. Changes in the level of arousal and also parity between levels were consistently reflected without impedance by treatment order. Inspection of the

data shows that treatment order produced variations in MSP but no order sequence was sufficiently strong to result in significant differences in MSP.

The question may be asked what accounts for the clear reflection of arousal in this band as compared to the other two bands investigated. The writer theorizes without giving substantiating proof that this level or range reflected arousal in general but not in emotion-specific terms. While some passages contained clear homosexual implications, others were heterosexual and again others pathological in nature. Passages resulting out of the five minute monologue contained a great variety of different emotions ranging from fear of shock-treatment, over loss of love object, fear of impotence, etc. to failure in university courses. The idea is tenable that the vocal patterns of different emotions have unique reflections in speech.

The reasoning of differences in vocal patterns, or emotion-specific patterns, has resulted in some interesting studies and has been substantiated to some extent. The work of Kramer which has been reported earlier is of interest here. Earlier studies in this area also supported Kramer's contention that listeners are able to judge specific emotions in standard passages which had been separated from emotion-specific introductions.

Compared to the NON-CLINICAL group, the CLINICAL subjects produced greater variations and subsequently more conflicting results in the narrow frequency band. If the existence of emotion-specific variations in the intensity-frequency relationship is tenable, it may be reasoned that the various passages produced conflicting arousal patterns during speech production in the emotionally more perceptive CLINICAL subjects. The high level of general arousal could have found its reflection in the wider 200 - 800 Hz band.

A less speculative reason for the differential results between the two groups, particularly for the narrow band, is the potential influence of medication in the CLINICAL group. It was generally observed that subjects belonging to this group suffered from a dried mouth during speech, and that they worked much harder to produce their speech sounds. It is worth pointing out, however, that regardless of differences in medication, degree of recovery, or psychiatric classification, the general arousal of these subjects was clearly reflected in the 200 - 800 band.

It has been reported, both by individuals having clinical experience with schizophrenics and in the research literature as well, that schizophrenics have distinct qualities in their voices that are distinguishable from the voices of non-schizophrenics (Goldfarb, Braunstein, and Lorge, 1956; Moskowitz, 1952; Spoerri, 1966). The secondary

hypotheses which were constructed to test for differences in Mean Speech Intensity between NON-CLINICAL and CLINICAL subjects were based on research which supported the above statement. The results of this study validate these earlier findings most positively. The difference in Mean Speech Intensity between the two groups of subjects were significant for all combinations of speech samples in each of the three frequency bands investigated. Without exception the Mean Speech Intensity of the CLINICAL group was lower than that of the NON-CLINICAL one.

The aforementioned findings bring to mind the work of Ostwald (1963) who in his discussion of voice prints compared those of people undergoing treatment with voice prints of 'normal' subjects. In Chapter II of this study, three voice prints (sharp, flat, and hollow), belonging to people in treatment, have been reproduced together with a fourth one which was produced by extroverted, confident people. A thought superimposition of voice prints 'flat' and 'hollow' over that of 'robust' would produce results which are germane to those reported in this study. Stating it in other words, normal, healthy persons produce more speech energy across the whole power spectrum than do those who suffer from depression over prolonged periods.

II Implications

The data obtained in this piece of research, particularly those obtained for the frequency bands 80 - 250 Hz and 200 - 800 Hz, would suggest that emotion-specific investigations adopting this or a modified technique of analysis could be fruitful. Some earlier research exists in this particular area which was conducted some thirty years ago by such researchers as Fairbanks, Pronovost, Hoaglin and others. These researchers presented listener-judges with a set of sentences which was common to five emotional passages spoken by experienced student actors. The listener-judges were asked to identify the emotion being represented in these sentences. By and large, these judges were successful in their performance but great variations were found between judges.

While the above studies would generate interesting insights into the expression of emotions in speech, a direct application of general changes in the level of arousal would be feasible in the experimental analysis of client-therapist communication during therapeutic interviews. The significance of this approach lies in the fact that, among other things, it permits objective measures of exceedingly subtle psychological phenomena or at least their precipitates; these may then be linked to clinical observations and constructs as well as to other autonomic devices. A study of

continuous speech using this or a modified approach would permit the analysis of variations of arousal within the individual interview as well as throughout the whole therapeutic relationship between therapist and client It seems theoretically possible to observe the impact of therapy on the client from the beginning to its conclusion in general terms, but in addition to this it would also allow for a more detailed analysis of client-specific problems by investigating specific speech responses as they recur during the process of therapy.

The observed difference in Mean Speech Intensity between the two groups of subjects seems to suggest fruitful applications in diagnostic work. The idea is tenable that with proper standardization of instruments and Mean Speech Intensity readings for various populations the former readings may be useful together with other indices in the diagnosing of pathological conditions in mental health work.

The writer is aware that certain crucial questions with regard to the recording of voices in a natural setting have to be solved in order to avoid the spilling of the voice of one partner into that of another. Limitations in variation from the recording microphone would present not only some technical problem but also some emotional ones for the subject studied as it would require him to wear a microphone

in a fixed position in such a way that neither head nor body movements would alter the set distance between mouth and microphone. There is no doubt in the writer's mind that advancing technology will answer the technical problems involved, but it will remain with the individual therapist to overcome the emotional ones.

III Summary

The present study has explored the feasibility of studying changes of arousal during speech within three frequency bands employing an acoustic approach in the analysis of sound pressure variations between sets of speech passages. Significant results in differences of arousal were obtained for one of the frequency bands analyzed in this piece of research. For the two groups of subjects studied, the bandwidth from 200 - 800 Hz consistently and accurately reflected changes in level of arousal or the parity between levels regardless of combination of passages employed, that is, whether the passages consisted of prepared neutral and emotional statements or of passages derived from the five minute monologue of dramatic events in the life of respective subjects. The narrower frequency band in comparison to the much wider one, 80 - 6300 Hz, produced also very good results which were acceptable for the NON-CLINICAL group of subjects, but it did not reflect

the same pattern for the CLINICAL group. A test on the difference in Mean Speech Intensity between the two groups of subjects produced statistically significant results with the CLINICAL group having the lower Mean Speech Intensity regardless of speech passage combination compared or bandwidth involved.

REFERENCES

REFERENCES

Addington, D.W. The Relationship of selected vocal characteristics to personality perception. Speech Monographs, 1968, 35, 492-503.

Allport, G., & Cantril, H. Judging personality from the voice. Journal of Social Psychology, 1934, 5, 37-55.

Alpert, M.S., Kurtzberg, R.L., & Friedhoff, A.J. Transient voice changes associated with emotional stimuli. Archives of General Psychiatry, 1963, 8, 362-365.

Aronson, E., & Mills, J., The effect of severity of initiation on liking for a group, Journal of Abnormal and Social Psychology, 1959, 59, 177-181.

Aronson, H. & Weintraub, W. Verbal productivity as a measure of change in affective status. Psychological Reports, 1967, 20, 483-487.

Assum, A.L., & Levy, S.J. Analysis of a nondirective case with follow-up interview. Journal of Social Psychology, 1948, 43, 78-79.

Auld, F., & Murray, E.J. Content-analysis studies of psychotherapy. Psychological Bulletin, 1955, 52, 377-395.

Bales, R F. Interaction Process Analysis. Cambridge: Addison-Wesley, 1950.

Barbara, D.A. Nonverbal communication. Journal of Communication, 1963, 13, 166-173.

Bernreuter, R. Personality Inventory. Stanford: Stanford-Univer, Press, 1931.

Boomer, D.S. & Dittmanm, A.T. Speech rate, filled pauses, and body movement in interviews. Journal of Nervous and Mental Disease, 1964, 139, 324-327.

Byrne, D., Anxiety and the experimental arousal of affiliation need, Journal of Abnormal and Social Psychology, 1961, 63, 660-662.

Byrne, D., & Sheffield, J., Response to sexual arousing stimuli as a function of repressing and sensitizing, Journal of Abnormal Psychology, 1959, 59, 177-181.

Campbell, D.T. Recommendations for APA test standards regarding construct trait, or discriminant validity. American Psychologist, 1960, 15, 546-553.

Cantril, H., & Allport, G. The Psychology of Radio. New York: Harper, 1935.

Dawes, R.M., & Kramer, E. A proximity analysis of vocally expressed emotions. Perceptual and Motor Skills, 1966, 22, 571-574.

Dibner, A.S. Ambiguity and anxiety, Journal of Consulting Psychology. 1958, 56, 165-174.

Diehl, C.J. Voice and personality; An evaluation. In D.A. Barbara (ed.), Psychological and psychiatric aspects of speech and hearing. Springfield: Charles C Thomas, Publisher, 1960, 171-203.

Dittman, A.T. The interpersonal process in psychotherapy: development of a research method. Journal of Abnormal and Social Psychology. 1952, 47, 236-244.

Dittman, A.T., & Wynne, L.C. Linguistic techniques and the analysis of emotionality in interviews. Journal of Abnormal and Social Psychology, 1961, 63, 201-204.

Dollard, J., & Mowrer, O.H. A method of measuring tension in written documents. Journal of Abnormal and Social Psychology, 1947, 42, 3-32.

Fairbanks, G., & Hoaglin, L.W. An experimental study of the durational characteristics of the voice during the expression of emotion. Speech Monograph, 1939, 6, 87-104.

Fay, P.J., & Middleton, W.C. Judgment of Spranger personality types from the voice as transmitted over a public address system. Character and Personality, 1939, 8, 144.

Fay, P.J., & Middleton, W.C. The ability to judge sociability from the voice as transmitted over a public address system. Journal of Social Psychology, 1941, 13, 303.

Fay, P.J., & Middleton, W.C. Judgment of introversion from the transcribed voice. Quart. Journal of Speech, 1942, 28, 226.

Feldstein, S., Brenner, M.S., & Jaffe, J. The effect of subject sex, verbal interaction, and topical focus on speech disruption. Language and Speech, 1963, 6, 229-239.

Fink, M., Jaffe, J., & Kahn, R.L. Drug induced changes in interview patterns: linguistic and neurophysiologic indices. In G.J. Sarwer-Foner (ed.) The Dynamics of Psychiatric Drug Therapy, Springfield, Ill.: Charles C Thomas, 1960, 29-37.

Fletcher, H. Speech and Hearing in Communication. New York: D.Van Nostrand Company, Inc., 1949.

French, N.R., & Steinberg, J.C. Factors governing the intelligibility of speech sounds. Journal of the Acoustical Society of America, 1947, XIX, 90-119.

Friedhoff, A.J., Alpert, M. & Kurtzberg, R.L. An effect of emotion on voice. Nature, 1962, 193, 357-358.

Friedhoff, A.J., Alpert, M. & Kurtzberg, R.L. Infra-content channels of vocal communication. In D.McK Rioch & E.A. Weinstein (Eds.) Disorders of Communication, 1964, 42, 414-423.

Gleser, G.C., Gottschalk, L.A., & Springer, K.J. An anxiety scale applicable to verbal samples. Archives of General Psychiatry, 1961, 5, 539-605.

Goldfarb, W., Braunstein, P. & Lorge, I. A study of speech patterns in a group of schizophrenic children. American Journal of Orthopsychiatry, 1956, 26, 544-555.

Gottschalk, L.A. Comparative psycholinguistic analysis of two psychotherapeutic interviews. New York: International University Press, Inc., 1961.

Gottschalk, L.A., Gleser, G.C., Magliocco, B.E. & D'Zmura, T. Further studies on the speech patterns of schizophrenic patients: Measuring inter-individual differences in relative degree of personal disorganization and social alienation. Journal of Nervous and Mental Disease, 1961, 132, 101-113.

Gottschalk, L.A., Kaplan, S.M., Gleser, G.C. and Winget, C.M. Variations in the magnitude of anxiety and hostility with phases of the menstrual cycle, Journal of Psychosomatic Medicine, 1962, 24, 300-311.

Gottschalk, L.A., Gleser, G.C. & Sringer, K.J. Three hostility scales applicable to verbal samples. Archives of General Psychiatry, 1963, 9, 254-279.

Gottschalk, L.A. & Frank, E.C. Estimating the magnitude of anxiety from speech. Behavioural Science, 1967, 12, 289-295.

Hargreaves, Wm.A., & Starkweather, J.A. Recognition of speaker identity. Language and Speech, 1963, 6, 63-67.

Hargreaves, Wm.A., Starkweather, J.A. & Blacker, K.H. Voice quality in depression. Journal of Abnormal Psychology, 1965, 70, 218-220.

Harris, R.E. Dominance, assertiveness, and hostility in persons with high and low blood pressure: Some similarities and some differences. Unpublished manuscript, University of California, at Berkeley, Institute of Personality Assessment and Research, 1953.

Holmgren, G.L. Physical and psychological correlates of speaker recognition. Journal of Speech and Hearing Research, 1967, 10, 57-66.

Jaffe, J. An objective study of communication in psychiatric interviews. Journal of the Hillside Hospital, 1957, 6, 207-215.

Jaffe, J., Language of the dyad: A method of interaction analysis in psychiatric interviews, Psychiatry, 1958, 21, 249-258.

Jaffe, J. Dyadic analysis of two psychotherapeutic interviews. In L.A. Gottschalk (Ed.) Comparative psycholinguistic analysis of two psychothereutic interviews. New York: International University Press, 1961, 73-90.

Jaffe, J., Fink, M., & Kahn, R.L. Changes in verbal transactions with induced altered brain function. Journal of Nervous and Mental Disease, 1960, 130, 235-239.

Kanfer, F.H. Verbal rate, eyeblink and content in structured psychiatric interviews. Journal of Abnormal and Social Psychology, 1960, 61, 341-347.

Kasl, S.V. & Mahl, G.F. Experimentally induced anxiety and speech disturbance. American Psychologist, 1958, 13, 349.

Kasl, S.V. & Mahl, F.G. The relationship of disturbances and hesitations in spontaneous speech to anxiety. Journal of Personality and Social Psychology, 1965, 1, 425-433.

Kauffman, P.E. An investigation of some psychological stimulus properties of speech behavior. Unpublished doctoral dissertation, University of Chicago, 1954.

Kersta, L.G., Voice print identification. Paper presented at the 63rd meetings of the Acoustical Society of America. New York. 1962/a.

Kersta, L.G. Voice print identification infallibility. Paper presented at the 64th meeting of the Acoustical Society of America. Seattle. 1962/b.

Kramer, E. Judgment of personal characteristics and emotions from nonverbal properties of speech. Psychological Bulletin, 1963, 60, 408-420.

Kramer, E. Elimination of verbal cues in judgments of emotion from voice. Journal of Abnormal and Social Psychology, 1964, 68, 390-396. (a)

Kramer, E. Personality sterotypes in voice: A reconsideration of the data. The Journal of Social Psychology, 1964, 62, 247-251. (b)

Krause, M.S. & Pilisuk, M. Anxiety in verbal behaviour: a validation study. Journal of Consulting Psychology, 1961, 25, 414-419.

Lasswell, H.D. Verbal references and physiological changes during the psychoanalytic interview: a preliminary communication. Psychoanalytic Review, 1935, 22, 10-24.

Licklider, J.C.R., & Miller, G.A. The perception of speech. In S.S. Stevens (Ed.), Handbook of experimental psychology, New York: Wiley, 1951, 1040-1074.

Mahl, G.F. Disturbances and silences in the patient's speech in psychotherapy. Journal of Abnormal and Social Psychology, 1956, 53, 1-15.

Mahl, G.F. On the use of "ah" in spontaneous speech: quantitative developmental, characterological, situational, and linguistic aspects. American Psychologist, 1958, 13, 349.

Mahl, G.F., Some observations about research on vocal behavior, In D. Mck. Rioch & E.A. Weinstein (Eds.), Disorders of Communication, 1964, 42, 466-483.

Mallory, E., & Miller, V. A possible basis for the association of voice characteristics and personality traits. Speech Monograph, 1958, 25, 255-260.

Markel, N.N., & Meisels, M. & Houck, J.E. Judging personality from voice quality. Journal of Abnormal and Social Psychology, 1964, 69, 458-463.

McReynolds, R. & Acker, M. On the assessment of anxiety. II: By a self-report inventory. Psychological Report, 1966, 19, 231-237.

Mehrabian, A, & Ferris, S.R. Inference of attitudes from nonverbal communication in two channels. Journal of Consulting Psychology, 1967, 31, 248-252.

Mehrabian, A., & Wiener, M. Decoding of inconsistent communications, Journal of Personality and Social Psychology, 1967, 6, 108-114.

Moore, W.E. Personality traits and voice quality deficiencies. Journal of Speech and Hearing Disorder, 1939, 4, 33-36.

Moses, P. The voice of neurosis. New York: Grune & Stratton, 1954.

Moses, P. Emotional causes of vocal pathology. In D.A. Barbara, Psychological and psychiatric aspects of speech and hearing, 1960, 204-231.

Moskowitz, E.W., Voice quality in the schizophrenic reaction type. Speech Monograph, 1952, 19, 118-119.

Ostwald, P.F. A method for the objective denotation of the sound of the human voice. Journal of Psychosomatic Research, 1960a, 4, 301-305. (a)

Ostwald, P.F. Visual denotation of human sounds. Archives of General Psychiatry, 1960, 3, 117-121. (b)

Ostwald, P.F. The sounds of emotional disturbance. Archives of General Psychiatry, 1961, 5, 587-592.

Ostwald, P.F. Soundmaking. Springfield, Ill.: Charles C Thomas, Publisher, 1963.

Ostwald, P.F. Acoustic methods in psychiatry. Scientific American, 1965, 212 (3), 82-91.

Overall, J.E. & Gorham, D.P. The brief psychiatric rating scale. Psychological Report, 1962, 10, 799-812.

Paivio, A. Personality and audience influence. In B.A. Maher (ed.), Progress in Experimental Personality Research, New York: Academic Press, 1965, 127-173.

Pittinger, R.E., Hockett, C.F., & Danehy, J.J. The first five minutes: A sample of microscopic interview analysis. Ithaca, N.Y.: Paul Martineau, 1960.

Pittinger, R.E., & Smith, H.L., Jr. A basis for some contributions of linguistics to psychiatry. Psychiatry, 1957, 20, 61-78.

Pope, B., & Siegman, A.W., The effect of therapist verbal activity and specificity on patient productivity and speech disturbances in the initial interview, Journal of Consulting Psychology, 1962, 24, 300-311.

Ramsay, R.W., Speech patterns and personality, Language and Speech, 1968, 2, 54-63.

Rubenstein, L., Electro-acoustical measurement of vocal responses to limited stress, Behavior Research and Therapy, 1966, 4, 135-138.

Sanford, F.H. Speech and personality. Psychological Bulletin, 1942, 39, 811-845.

Shapiro, J.G. Variability in the communication of affect. The Journal of Social Psychology, 1968, 76, 181-188.

Siegman, A.W. & Pope, B. Effects of question specificity and anxiety-producing messages on verbal fluency in the initial interview. Journal of Personality and Social Psychology, 1965, 4, 522-530.

Skinner, E.R. A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness. Speech Monograph, 1935, 2, 81-137.

Smith, J.E. Keith. Decision-theoretic speaker recognizer. Paper presented at the 64th meetings of the Acoustical Society of America (Seattle).

Soskin, W.F. Some aspects of communication and interpretation in psychotherapy. Paper read at American Psychological Association, Cleveland, September 1953.

Soskin, W.F. & Kauffman, P. Judgment of emotion in word-free voice samples. Journal of Communication, 1961, 11, 73-80.

Spoerri, T.H. Speaking Voice of the schizophrenic patient. Archives of General Psychiatry, 1966,14.

Starkweather, J.A. Judgments of content-free speech as related to some aspects of personality. Unpublished doctoral dissertation, Northwestern University, 1955.

Starkweather, J.A. Content-free speech as a source of information about the speaker. Journal of Abnormal and Social Psychology, 1956, 52, 394-402, (a)

Starkweather, J.A. The communication-value of content-free speech. American Journal of Psychology, 1956, 69, 121-123, (b)

Starkweather, J.A. Vocal communication of personality and human feelings. Journal of Communication, 1961, 11, 63-86.

Strupp, H.H. Psychotherapists in Action, New York: Grune & Stratton, 1960.

Snyder, W.U. An investigation of the nature of non-directive psychotherapy. Journal of General Psychology, 1945, 33, 193-233.

Snyder, W.U. A comparison of one unsuccessful with four successful non-directively counseled cases, Journal of Consulting Psychology, 1947, 11, 38-42.

Sullivan, H.S. The psychiatric interview. New York: Norton, 1954.

Szedenik, A.M., Psycho-sexual Problems, Derby, Conn.: The New International Library, Inc., 1964.

Thompson, C.W. & Bradway, K. The teaching of psychotherapy through content-free interviews. Journal of Consulting Psychology, 1950, 14, 321-323.

Trager, G.L. Paralanguage: A first approximation. Stud. Linguist., 1958, 13, 1-12.

Voiers, W.D. Perceptual bases of speaker identity The Journal of the Acoustical Society of America, 1964, 36, 1065-1073.

Weintraub, W. & Aronson, H. The application of verbal behavior analysis to the study of psychological defense mechanisms: Methodology and preliminary report. Journal of Nervous and Mental Disease, 1962, 134, 169-181.

Weintraub, W. & Aronson, H. The application of verbal behavior analysis to the study of psychological defense mechanisms. II: Speech pattern associated with impulsive behavior. Journal of Nervous and Mental Behavior, 1964, 139, 75-82.

Weintraub, W. & Aronson, H. The application of verbal behavior analysis to the study of psychological defense mechanisms. Journal of Nervous and Mental Disease, 1965, 141, 172-179.

Weintraub, W. & Aronson, H. The application of verbal behavior analysis to the study of psychological defense mechanisms. IV: Speech pattern associated with depressive behavior. Journal of Nervous and Mental Disease, 1967, 144, 22-28.

Winer, B.J., Statistical Principles in Experimental Design, New York: McGraw-Hill Book Company, 1962.

Zimbardo, P.G., Mahl, G.F. & Barnard, J.W. The measurement of speech disturbance in anxious children. Journal of Speech and Hearing Disorders, 1963, 28, 362-370.

APPENDIX A

Analyzed Voice Samples

Analyzed Voice Samples

Neutral Speech Samples (NSS)

The men work the year round as laborers on the farm and have their wages paid at the end of the month during the summer, but during the winter time when harvesting and processing have come to an end, payment is made less frequently.

And he proceeded to make the molten sea ten cubits from its one brim to its other brim, circular all around, and its height was five cubits and it took a line of thirty cubits to circle all around it when it was finally built after many months of work.

Change is a circular process that proceeds from the formal to the informal to technical to new formal with the emphasis shifting rather rapidly at certain junctures which mark the breaking point between the systems that have been formed.

Thus the stomach is compressed and its contents being prevented from passing downward by the firm contraction of the pyloric region are forced through the relaxed cardia on to the esophagus which begins to relax along its total length.

The tongue is a movable organ that performs the important functions concerned with taste, mastication, swallowing, and speech and is therefore composed of muscles and covered with a mucuous membrane to perform these functions adequately.

On the open range with cattle belonging to the different ranches an annual round-up is a necessity in order to brand and count the cattle belonging to each of the ranches bordering on the rich and fertile grasslands.

Emotional Speech Samples (ESS)

He lifted up my shirt and with his busy fingers fell to visit and explore that part of me where now the heat and irritation were so violent that I was perfectly sick and ready to die with desire, which his sensitive touch increased beyond comprehension.

Yet my hand knew, too, how to unclothe her where I wanted it, and then I drew down the thin silk sheath, slowly, carefully, right down and over her feet, and with a quiver of exquisite pleasure I touched the warm soft body, and touched her navel for a moment in a kiss.

I laid my hand on her shoulder, and softly, gently, it began to travel down the curve of her back, blindly, with a blind stroking motion, to the curve of her crouching loins, and there my hand, softly, softly, stroked the curve of her flank, in the blind instinctive caress.

I drew his breeches quite down to his knees, and while I approached he still kept working and grinding his belly against the cushion under him between which I had insinuated my hand softly touching his thigh and reaching for the small of the back more firmly.

I usually went at dusk and secreted myself close to a women's toilet where I could hear and smell the excretory process; I felt a strong desire and high sexual excitement and usually masturbated while urinating and sometimes I even had an emission without masturbation.

I stayed firm inside her, given to her, while she was active, passionately active, coming to her own crisis, and as I felt the frenzy of her achieving her own orgasmic satisfaction, from my hard, erect passivity, I had a curious sense of pride and satisfaction.

Anxiety Speech Samples (ASS)

The following samples have been chosen from the SSS of different subjects. The score on the left of each passage indicates type of anxiety experienced and its respective weight.

2a3 there was the wolf sitting down the other side of it/
6a3 looking up at me/
2a3 all in a in a crouched position/
2a3 like it was ready to leap/
1a3 and so this just about stopped my heart right on the spot/

4a3 how attached you get to it/
3b2 and when she was put away/
3a3 you had that emptiness inside of you/
4a3 just wondering why you ever/
4a3 why this would have been allowed/

5a3 when my first wife chased around a bit/
5a3 caused a bit of a problem/
3a3 and I ended up in divorce court/
5a3 I dare not say any more/

3a3 and I did remember having to go back again/
3a3 I went back down to Manitoba/
3a4 and I was very unhappy there/
3a3 leaving my parents and going down/
3a3 and how hard it was for me to go back there/

Anxiety-free Speech Samples (AfSS)

but ah about a mile from the mine/
there was at the base of a mountain a hot spring/
where the previous people there had dug a hole in the gravel/
and would bathe in it/
there was water about hundred degrees/

I went to university in the fall/
and I was registered in a seminary/
which was what my brother was in/
and I had decided/
I was going to go into seminary/

and the sun was just starting to rise when we left/
it was about seven o'clock in the morning/
and we decided to drive out to C./
let's see it was the second, I imagine/
and it started off pretty mundane/
I guess it was fairly cold/

and a friend of mine and his girlfriend came down about 12:30/
so we had coffee at our place for a while/
and we decided to go to a party with them/
so we all got into the car/
and we started to go/

APPENDIX B

Recording Set-Up

APPENDIX C

Technical Specifications of Equipment

Technical Specifications of Equipment

AMPEX AG-500 Recorder/Reproducer

Overall Frequency Response:	15 ips:	± 2 dB, 30 Hz - 18 kHz
	7½ ips:	+ 2 dB - 4 dB, 30 Hz - 15 kHz
	3 3/4 ips:	+ 2 dB - 4 dB, 40 Hz - 8 kHz
Signal to Noise Ratio:	55 dB at 15 and 7½ ips (half or two track) 60 dB at 15 and 7½ ips (full track) 55 dB at 3 3/4 ips (full track)	
Flutter and Wow:	Less than 0.15% rms at 15 ips 0.18% rms at 7½ ips 0.25% rms at 3 3/4 ips	

AMPEX AV-770 Recorder/Reproducer

Overall Frequency Response:	7½ ips:	± 3 dB, 50 Hz - 15 kHz
	3 3/4 ips:	± 3 dB, 50 Hz - 8 kHz
Signal to Noise Ratio:	46 dB at 7½ ips 43 dB at 3 3/4 ips	
Flutter and Wow:	0.15% rms at 7½ ips 0.20% rms at 3 3/4 ips	

HEWLETT-PACKARD Model 2010J Data Acquisition System (DYMEC)

The Data Acquisition System measures analog data derived from a number of sources, and displays and records this information in digital form. It uses as Integrating Digital Voltmeter as digitizer and a guarded Crossbar Scanner. Range of Frequency Measurements 5 Hz to 300 kHz; Voltage range from 0.1 V to 1000 V full scale with polarity sensed and indicated automatically.

HEWLETT-PACKARD Model 3400A RMS Voltmeter

Voltage range: 1 mV to 300 V, 12 ranges; dB range: -72 to +52 dBm; Frequency range: 10 Hz to 10 MHz; Response: responds to rms value of the input signal for all waveforms; AC-to-DC converter accuracy: ± 5% over the range from 10 Hz to 10 MHz.

SKL - Variable Electronic Filter, Model 308A

Cut-off Frequency Range: 0.2 Hz to 20 kHz in five decade ranges; Accuracy of cut-off frequency: ± 3.5%; Attenuation at cut-off frequency: 3.0 dB; Rate of Attenuation in Rejection Band: 24 dB per octave per section; Maximum attenuation: greater than 80 dB; Insertion loss: 4.5 ± 1.0 dB.

Diagrammatical Sketch of Circuitry