

AN ANALYSIS OF VOICEPRINT
IDENTIFICATION

Thesis for the Degree of M. S.
MICHIGAN STATE UNIVERSITY

JAMES J. HENNESSY

1970

AN ANALYSIS OF VOICEPRINT

IDENTIFICATION

By

James J. Hennessy

AN ABSTRACT OF A THESIS

Submitted to the College of Social Science
Michigan State University
in partial fulfillment of the requirements
for the Degree of

MASTER OF SCIENCE

School of Criminal
Justice

1970

APPROVED

Robert C. Tegenau
(Chairman)

James H. Conroy
(Member)

Robert F. Turner
(Member)

ABSTRACT

AN ANALYSIS OF VOICEPRINT IDENTIFICATION

By James J. Hennessy

The purpose of this thesis is to investigate the validity, reliability, and feasibility of voiceprint identification for use by law enforcement agencies in the investigation of crimes involving speech communication and the identification of suspects by their voices. To analyze the validity, reliability, and feasibility, an experiment was carried out. The subsequent results indicated that, while voiceprint identification may be feasible, the actual determination of its reliability and validity awaits further research. The research data of the opponents of the technique of voiceprint identification conflicts with the experimental results of those who favor the technique. The experimental data obtained from the research of the School of Criminal Justice reported in this thesis was not intended to definitively answer the questions of the validity and reliability of voiceprint identification: nor has it done so.

The future tests of voiceprint identification will hopefully determine the scientific and legal value of the voiceprint technique and end the present dispute concerning the reliability, validity, and feasibility of voiceprint identification.

AN ANALYSIS OF VOICEPRINT
IDENTIFICATION

By

James J. Hennessy

A THESIS

Submitted to the College of Social Science
Michigan State University
in partial fulfillment of the requirements
for the Degree of

MASTER OF SCIENCE

School of Criminal
Justice

1970

G-65682
1-27-71

ACKNOWLEDGMENT

Grateful thanks are expressed to Mr. Clarence H.A. Romig and Mr. Vernon Rich as well as the other members of my thesis committee for their assistance, encouragement, and direction in carrying out this research and guidance in the preparation of this thesis.

Appreciation must also be expressed to Dr. Oscar Tosi and Dr. Carl Pedrey of the Audiology and Speech Sciences Department of Michigan State University for their aid in this project.

Sargeant Ernest Nash and Trooper Louis Wilson must also receive my sincere gratitude for their technical assistance and encouragement.

Lasting gratitude is due to Mr. James A. Keyes and Mr. David Metzger without whom this research could not have been carried out.

Lastly, my deepest appreciation is expressed to my wife, Sally, who encouraged me and aided me, and whose patience knows no end.

TABLE OF CONTENTS

	Page
List of Tables	vi
 <u>Chapter</u>	
I. INTRODUCTION	1
THE PROBLEM	2
Statement	2
Purpose	3
Scope	4
Importance	4
ORGANIZATION OF THE THESIS	5
DEFINITIONS	7
HISTORY	8
II. SOUND, SPEECH, PHONETICS, AND VOICEPRINT IDENTIFICATION	12
INTRODUCTION	12
SOUND	13
Resonance	14
Sound Waves	16
SOUND AND SPEECH	17
Speech Processes	17
Modulation	18
From Sound to Speech	20

	Page
PHONETICS	21
Sounds of a Language: Phonemes and Phones . .	21
Types of Sounds	22
Vowels	23
Diphthongs	23
Semivowels	24
Consonants	24
Glides	25
Nasals	25
Laterals	25
INVARIANT SPEECH	26
Sound Changes	27
Language Changes	28
Speaker Articulation Changes	28
Transitions Between Sounds	29
Corollary of the Theory of Invariant Speech	30
SOUND SPECTROGRAPH	30
Function	30
Sound Energy Analyzed	32
Results of Analysis	34
CHAPTER SUMMARY	35

Chapter		Page
III.	A REVIEW OF THE LITERATURE CONCERNING EXPERIMENTS WITH VOICEPRINT IDENTIFICATION	36
	AURAL RECOGNITION EXPERIMENTS	37
	General Characteristics	37
	On Identification of Speakers by Voice	38
	Perceptual Basis of Speaker Identity	39
	Effects of Stimulus Content and Duration on Talker Identification	41
	Section Summary	42
	SPECTROGRAPHIC ANALYSIS	43
	Kersta's Experiments	43
	First Experiment	43
	Second Experiment	45
	Third Experiment	46
	Fourth Experiment	47
	Miscellaneous Tests	47
	Section Summary	49
	CONFLICTING DATA	49
	Effects of Context on Talker Identification	50
	Speaker Authentication and Identification	55
	CHAPTER SUMMARY	57

Chapter		Page
IV.	RESEARCH OF THE SCHOOL OF CRIMINAL JUSTICE OF MICHIGAN STATE UNIVERSITY	59
	THE RESEARCH PROJECTS	59
	Goals	59
	Planning	60
	Identifiers	62
	Training	62
	Pilot Study	65
	Procedures	66
	Speech Samples and Speakers	67
	Recording Tapes and Spectrograms	68
	Identification Tasks	68
	Results	70
	PRINCIPAL STUDY	71
	Factors Under Study	71
	Procedures	73
	Analysis of Recordings	76
	Results	77
	Audiology and Speech Sciences' Results	81
	CHAPTER SUMMARY	82
V.	THE LEGAL DILEMMA OF VOICEPRINT	83
	VOICEPRINT IN COURT	84
	United States v Wright	84
	People v Straehle	84
	California v King	85
	New Jersey v Di Gilio	86
	New Jersey v Cary	87

Chapter	Page
GENERAL ADMISSIBILITY REQUIREMENTS	89
The Frye Rule	89
VOICEPRINT IDENTIFICATION AND CONSTITUTIONAL LAW . .	92
Fourth Amendment	92
Fifth Amendment	93
Sixth Amendment	94
CHAPTER SUMMARY	95
VI. SUMMARY AND CONCLUSIONS	98
CONCLUSIONS	98
PROPOSALS	99
BIBLIOGRAPHY	101
APPENDIX A. Specifications of the Voiceprint Laboratories's Sound Spectrograph	113
APPENDIX B. Specifications of Tape Recorders Used	114
APPENDIX C. Frequency Response of Microphones Used	115

LIST OF TABLES

Table	Page
1. Results of Bricker and Pruzansky's 1966 Experiment . . .	42
2. Results of Kersta's 1962 Experiment	45
3. Results of Cambell and Young's 1967 Experiment	53
4. Results of the 1969-1970 Pilot Study of The School of Criminal Justice	70
5. Results of the Chi Square Statistical Analysis of the Pilot Study Data of the School of Criminal Justice	70
6. Results of the Main Study of the School of Criminal Justice	73
7. Results of the Chi Square Statistical Analysis of the Main Study Data of the School of Criminal Justice . .	79

Chapter 1

INTRODUCTION

One of the underpinnings of civilization is the communication of information. Speech is a basic manner of communication. It would be difficult, if not ridiculous, to think of the development of a civilization without the utilization of speech. And, today, even with computerization and teleprinting, speech remains a fundamental form of the communication of information.

There are certain acts involving speech communication, however, which, because of laws, are crimes. Bomb threats, extortion demands, false fire alarm calls, kidnap-ransom demands, and obscene telephone calls are all acts using speech as the primary means of communication of information. These crimes are extremely difficult to prevent or solve. The caller is usually anonymous, and a pay telephone is often used. Prior to 1962, any attempt to identify the caller in an obscene telephone call, for instance, was limited to the victim's recognition of the caller's voice or the tracing of the telephone call. A memorable case in point is the Hauptmann case. Hauptmann's voice was identified by Charles Lindbergh as that of his child's kidnapper. The identification took place some 2 years after Lindbergh had heard the kidnapper's voice. The chances of actually convicting a suspect were even less than identifying him in the first place.

¹"Break in the Greatest Story of Newspaper History," Newsweek 5 (February 23, 1935), 23-24.

1

However, speech does more than convey just the information carried in the words spoken. By listening to a suspect's voice and the recorded voice of a caller, a victim of an obscene telephone call may be able to identify the suspect or reject him. In 1962, the technique of acoustic spectrography was unveiled by Mr. Lawrence G. Kersta. He claimed that voiceprint identification, his term for the technique of acoustic spectrography, had reached a level of development which made possible positive identifications of suspects from their voices. Kersta stated that through voice analysis by a sound spectrograph and visual analysis and comparison of the spectrograms, a person could be identified by his voice. From Kersta's initial claims to the present time, a controversy has surrounded voiceprint identification. Kersta claims the technique is infallible; Peter Ladefoged and Ralph Vanderslice claim it is a farce.²

THE PROBLEM

Statement of the Problem

The very fact that there is such a great amount of controversy about voiceprint identification raises the question of its real worth for law enforcement purposes in investigation of crimes and identification of suspects. Kersta has carried out experiments with results purporting to show that voiceprint is 99.7% accurate.³ Peter Ladefoged,

²Dr. Oscar Tosi, personal interview, August 17, 1970; see also Lawrence G. Kersta, "Voiceprint Infallibility" (unpublished paper presented to the Acoustic Society of America, November 7, 1962); also see M. Young and R. Cambell, "Effects of Context on Talker Identification," Journal of the Acoustical Society of America, 42 (1967), 1250-56; and Carl Asp, "Voiceprint: Its Evolution, Application and Present Status" (unpublished paper, University of Tennessee).

³Lawrence G. Kersta, "Voiceprint Identification", Nature, 196 (December 29, 1962), 1253-57.

an expert in phonetics, denies the validity and reliability of the voiceprint technique; and results have been obtained showing that the technique has little better than chance accuracy (51%).⁴

The problem remains: can a person be positively identified by his voice through the technique of acoustic spectrography? The implications of the answer to this question are important. If the technique is feasible, voiceprint would be an invaluable aid for the investigation of crimes involving speech communication and the identification of suspects by their voices. In addition, voiceprint identification must be guarded and restricted in order to prevent its manipulation for repressive and unconstitutional purposes or its degeneration into a "hocus-pocus" method. On the other hand, if the technique is invalid and unreliable, its use as a tool for positive identification must be stopped; its use as an investigative method must be marked with extreme caution.

Purpose

It is the purpose of this study (1) to examine the technique of acoustic spectrography and its theoretical bases; (2) to review and analyze experiments previously performed concerning the technique of acoustic spectrography; (3) to submit a new experiment and the results for examination; (4) and to investigate the relationship of voiceprint identification to the general requirements for its admissibility as evidence in court and the requirements provided by the Fourth, Fifth, and Sixth Amendments of the Constitution of the United States.

⁴K. Stevens and others, "Speaker Authentication and Identification: A Comparison of Spectrographic and Auditory Presentations of Speech Material," Journal of the Acoustical Society of America, 44 (1968), 1596-1607.

Scope

The scope of this research study encompasses four non-technical areas of the acoustic spectrograph and the technique of acoustic spectrography. (1) The theories, acoustic, phonetic, biological, and physical, upon which the technique is based are considered. The actual electrical and mechanical functioning of the sound spectrograph is not considered to be within the scope of this thesis. (2) The reliability and validity of the sound spectrograph and the technique of acoustic spectrography in relation to the uses of these techniques by law enforcement are examined in the light of the experimental data available. The experiments and arguments of the proponents as well as those of the opponents are presented. (3) The research experiments of the School of Criminal Justice of Michigan State University are reported and discussed. The research project of the Audiology and Speech Sciences Department of Michigan State University is also briefly discussed. (4) Finally, some of the legal aspects of voiceprint identification are considered.

Importance of the Study

Experimental confusion and conflict as well as precipitant use of the technique before it was scientifically substantiated have resulted in the present controversy over voiceprint identification. An experimentally objective evaluation of the technique of acoustic spectrography as used for law enforcement purposes is the aim of this study. The importance of the study rests on this goal.

ORGANIZATION OF THE REMAINDER OF THE THESIS

Chapter 1

The contents of this thesis have been divided into six chapters. The preceding portion of Chapter 1 has introduced the thesis subject, stated the problem, delineated the scope of this thesis, and asserted the importance of the problem. The remaining portion of this Chapter contains, one, definitions of terms used in the thesis which may be either unfamiliar to the reader or inadequately defined for the requirements of clarity needed in this subject; and, two, a brief history of the development of visible speech and the technique of acoustic spectrography is presented. The history is concise and considers only those points necessary for a general understanding of the original purpose of the development of the acoustic spectrograph and the outgrowth of the technique of acoustic spectrography, voiceprint identification.

Chapter 2

Chapter 2 contains a discussion and explanation of the theories upon which voiceprint identification is based. One of the major problems with voiceprint identification is that its theoretical foundations cut across the fields of acoustics, phonetics, physics, biology, and law. A succinct examination of the contributions of these fields is presented.

Chapter 3

The issues of the controversy generated by voiceprint identification in its use by law enforcement are illustrated in this chapter. The chapter is divided into the case for and the case against voiceprint

identification. The literature will be reviewed in this chapter in order that both factions' views may be presented.

Chapter 4

A report of the experimental research of the School of Criminal Justice comprises the fourth chapter. It presents an original study aimed at helping to clarify the feasibility of voiceprint identification. The planning, training, pilot study, design of the main study, methodology involved in the execution of the design, the results, and a discussion of the results are contained in this chapter. A brief discussion of the results of the research project of the Audiology and Speech Sciences Department is also contained in the chapter.

Chapter 5

In Chapter 5, the legal aspects of voiceprint identification are examined. The cases in which voiceprint identification has been used are presented with their results. The basic question of whether or not voiceprint identification is admissible evidence is discussed. The basic requirements of the Fourth, Fifth, and Sixth Amendments with regard to the voiceprint technique are presented with reference to the most recent cases applicable to voiceprint identification.

Chapter 6

A summary and conclusions of the materials and data presented in the preceding chapters compose the final chapter of the thesis. An objective evaluation of the research, data, and practical applications as reported in the thesis is presented. Partial conclusions concerning the reliability, validity, and feasibility of voiceprint identification are also included. Proposals as to further research and proper use by law enforcement officers are developed.

DEFINITIONS

The following are terms that are used throughout the thesis and which must be clearly comprehended if the topic is to be understood.

Acoustic Spectrography

Acoustic Spectrography is the analysis of sound by a sound spectrograph. The sound spectrograph analyzes sound, using the parameters of frequency, intensity, and duration.

Kay Sonagraph

The Kay Sonagraph is the sound spectrograph made by the Kay Electric Company of New Jersey. It is similar to the Voiceprint Laboratories' sound spectrograph, also made in New Jersey.

Sound Spectrogram

The Sound Spectrogram is the permanent visual record produced by the sound wave analyzer, the sound spectrograph.

Sound Spectrograph

The Sound Spectrograph is a sound wave analyzer which produces a permanent visual record showing the distribution of sound energy in frequency, time, and intensity.

Technique of Acoustic Spectrography

This term is more technically correct than the term, voiceprint identification. The latter is a popular and somewhat inaccurate description of the technique of acoustic spectrography. The technique of acoustic spectrography itself may be defined as the subjective comparison of the visual results of an analysis of voices by a sound spectrograph.

Voiceprint

A voiceprint is the sound spectrogram of a voice.

Voiceprint Identification

Voiceprint identification is the popular term, coined by Mr. Lawrence G. Kersta, for the technique of acoustic spectrography.

A CONCISE HISTORY OF ACOUSTIC SPECTROGRAPHY

During the Second World War, the United States Government requested Bell Telephone Laboratories in New Jersey to investigate the practicality of a technique to identify an individual by his voice. The purpose of such a technique was that, if an individual could be rapidly identified by his voice, the United States forces could identify German radiomen and those individuals whose radio transmissions were intercepted by the United States forces. The ultimate aim was to follow troop movements by following the radiomen attached to the various troop groups. For example, if one radioman were recorded in one sector at one time and later recorded in another sector, it might be concluded that his company had moved into a new sector.⁵

Bell, in 1941, accepted the Government's proposal and put five scientists, including Kersta, on the project.⁶ Although the scientists made some progress toward accomplishing their goal, they were able to

⁵Sargeant Ernest Nash, (unpublished paper), 1.

⁶Harry More, "Voiceprint - A New Scientific Aid for Law Enforcement," National Sheriff (March-April, 1969), 12.

develop neither an instrument nor a technique with a useful degree of efficiency and reliability.⁷

However, from their research, the concept of visible speech took form. In 1943, a group of normal girls under the instruction of Ralph Potter, George Kopp, and Harriet Green, researchers at Bell, learned to read the spectrograms made from the spectrographic analysis of human voices. They could read both the stationary and moving spectrographic patterns. What this meant was that ongoing speech was able to be analyzed by a sound spectrograph, either projected on a screen or made into spectrograms, and read and understood by a person who had never heard the words spoken. It had originally been the hope of these investigators to make the use of the telephone available through this means to the totally deaf.

To this end, the experimenters created a direct translator. The direct translator was the same as the sound spectrograph except that the patterns of speech were projected on a screen rather than made into spectrograms. The girls were trained to be able to identify the sounds spoken, and, literally, to read the ongoing speech.

There were two types of direct translators. The first used twelve fixed filters, each with a bandwidth of 300 cycles per second (c.p.s.; now called hertz, Hz). Each filter analyzed the frequencies within its width. The output of each filter appeared as a trace of light on a moving screen of phosphorescent material. The dimensions of frequency, intensity, and time were represented. This first translator proved to be too large, and a second, smaller machine was built. This model had

⁷Nash, op. cit., 2.

a rotating drum of phosphorescent material and twelve filaments. Each filament was connected to a filter whose output was traced on the rotating drum.⁸

The experimenters hoped that their research would prove that their technique was not only possible but of great usefulness to the deaf. The amount of time and the complexity of the training, as well as the cost of the machine, however, contributed to the failure of the concept to gain widespread use.

Kersta, meanwhile, had continued his research on the sound spectrograph. Bell had received a request from the New York City Police Department for aid in tracking down bomb-scare callers.

Finally, Kersta requested and was granted by Bell a leave of absence to devote his full time to the sound spectrograph.⁹ In 1962, he wrote an article in Nature magazine describing his results.¹⁰ He applied to Bell Telephone Laboratories for a license to manufacture and market his machine and technique. Bell granted the license.

Since then, Mr. Kersta has used his machine and method for state and local agencies of law enforcement as well as for the United States Air Force, Civil Aeronautics Board, Federal Aviation Agency, and the Defense Department.¹¹ In his work, he has unscrambled last messages from airliners that mysteriously crashed,¹² and identified one of the

⁸Ralph Potter, George Kopp, and Harriet Kopp, Visible Speech (New York, New York: Dover Publications Incorporation, 1947), p. 5.

⁹Nash, op. cit., 2.

¹⁰Kersta, "Voiceprint Identification," op. cit., 1253-57.

¹¹More, op. cit., 26.

¹²Ibid., 26.

participants in a secret communication involving a plan to blame the United States and Great Britain for the Arab defeat in the 1967 Seven Days War.¹³ Kersta might have identified the other voice, probably that of King Hussein of Jordan, if a known sample of Hussein's voice had been available. Most recently, Kersta's technique has been used by law enforcement agencies for the identification of suspects by their voice. He has testified in numerous cases with mixed results. Much of his testimony has been contradicted by other experts who do not believe his technique actually works.

This history brings us to the present. The controversy surrounding voiceprint identification still exists.

¹³Peter Ladefoged and D. Broadbent, "Information Conveyed by Vowels," Journal of the Acoustical Society of America, 29 (1957), 66.

Chapter II

SOUND, SPEECH, PHONETICS, AND VOICEPRINT IDENTIFICATION

INTRODUCTION

If voiceprint were reliable and valid, then it would present a new and invaluable tool for law enforcement in preventing crime and apprehending criminals.

"In the first six months of 1969, for example, the Bell System verified the placing of 306,103 obscene or harassing telephone calls. But tens of thousands of others go unreported.

"In the same period, there were 18,115 threatening calls.

"Such a call makes a person feel helpless. For one thing, the caller is anonymous. For another, there is no way of knowing when the phone rings whether the caller is a family member, friend, or on legitimate business, or whether it is some freak."¹⁴

The technique of acoustic spectrography could be one practical means of not only apprehending "freaks," but also of deterring such crimes as these. The technique, however, needs to be verified first.

One of the problems of the voiceprint controversy stems from the question of who should judge, who should determine whether or not the technique is reliable and valid. The question is asked because the

¹⁴Bernard Gavzer, "On Guard," State Journal (Lansing), September 25, 1970, p. A 12.

theoretic and practical framework of voiceprint identification draws on four main sources: acoustics, physics, biology, and phonetics.

In order to understand the basic hypotheses of the technique of acoustic spectrography, it is necessary to have an understanding of the contributions of the four sources. A brief but adequate examination of the amassed information drawn from these four areas is presented.

One point must be made clear before the various data sources are investigated. The viewpoint of the examiner will determine the terms used and the definitions of those terms. For example, if a psychological viewpoint is used, the term for the strength of a sound is loudness. Loudness is partially defined by the perception of the sound by the ear. On the other hand, if a different view is taken, that of physics, then the term loudness becomes amplitude, and the perception of the sound by the ear is in no way part of the definition of the term amplitude. For purposes of consistency and intelligibility, the terms of the science of physics are used in this study.

SOUND

Speech is sound. To understand speech, one must understand something about sound.

"Sound is a condition of disturbance of the particles of an elastic medium which is propagated in a wave outward in all directions from a vibrating body and takes the form of the displacement of the particles forward and backward from their position of rest in the direction of the propagation of the sound."⁴⁵

⁴⁵Richard Hoops, Acoustics in Speech (Springfield, Illinois: Thomas Publisher, 1960), p. 5.

Sound is basically the back and forth vibration of molecules of air. When, for example, one molecule is set in motion, it collides with those molecules around it, setting these in motion. As the particles are propelled forward by these collisions, a rarefaction, a vacuum, develops behind the outward moving particles. The neighboring particles continue to be pushed on, however, and cause the compression of molecules to progress in the direction in which the sound is traveling. The distance travelled by the molecules is the amplitude of the sound and is characterized by the height of a sine wave.¹⁶ The particles are finally sucked back into the rarefaction which their movement forward created behind them. A cycle is the movement of molecules forward and backward. The number of movements forward and backward in a measured period of time is the frequency and is characterized by the number of sine waves along the time axis of an oscillogram.¹⁷

Resonance

The successive collisions of the molecules produce sound waves. When the waves strike an unyielding substance, an echo is produced. However, if the substance will vibrate at the frequency of the sound waves, and if the waves strike it with enough force, the substance will resonate and produce its own sound waves at the same frequency. The standard example is the relationship of two tuning forks which vibrate at the same frequency. If one of the two forks, placed beside the other, is struck, the second will also begin to vibrate. However, if

¹⁶Bernard Kamine, "The Voiceprint Technique of Speaker Identification: Its Validity and Admissibility in Court," (unpublished paper, March, 1968), 4.

¹⁷Ibid., 4.

four forks are placed side by side and two are of the same frequency and two are not, and, again, if one of the two with the similar frequency is struck, the other fork of the same frequency will vibrate, but the other two forks of dissimilar frequencies will not resonate. Sympathetic vibration occurs when the resonator is in tune, of the same frequency, with the generator.¹⁸ The period at which a vibrating body resonates with the greatest ease is known as its natural period.¹⁹

The vibrations or resonance created by a generator upon a resonator are of two types, periodic and aperiodic. Periodic vibrations are vibrations in a regular pattern; the vibrations repeat themselves in equal intervals of time. This type of vibration is also called harmonic motion. Music is harmonic motion. Aperiodic vibrations are vibrations in erratic or irregular patterns; they do not repeat themselves in equal intervals of time. Speech is a good example of aperiodic vibration.²⁰

For resonance to occur, there must be a source of energy, a generator, a substance capable of being vibrated, a resonator, and a transmitting medium, air.

One of the characteristics of resonance is the relationship of the amplitude of the resonance to the frequency of the resonator. The more similar the frequency of a resonator is to the frequency of the source of the energy, the greater the amplitude of the vibrations; the more dissimilar, the less the amplitude is. In addition, if a resonator is bombarded with many frequencies, the resonator will act as a

¹⁸Lyman Judson and Andrew Weaver, Speech Science (New York, New York: Appleton-Century-Crofts, 1965), p. 91.

¹⁹Lyman, Ibid., p. 89.

²⁰Hoops, op. cit., p. 6.

selective transmitter. The frequencies most similar to the frequency of the resonator will be selected or passed to resonate; the frequencies not as similar or completely dissimilar will not be passed.²¹ The resonator, therefore, may amplify, damp, or destroy vibrations.

Natural resonators, as distinguished from mechanical ones, are the resonators of particular interest for this study. Natural resonators are of two types, sounding boards and cavities. The selectivity of a sounding board is fixed by its size, mass, and texture. The selectivity of a cavity may be fixed or variable if its dimensions, openings, and textures are variable.

Sound Waves

Sound forms a pattern of molecules of air moving forward and backward. This pattern is called a sound wave.²² One of the simplest kinds of sound waves with a regular pattern is the pure tone. The pure tone is a sound of constant frequency and amplitude. It has a wave form with the shape of a sine wave.²³ For practical purposes, though, the most common wave form is the complex wave. Two pure tones can combine to form a complex wave pattern. Speech is made up of complex wave forms.

²¹Claude Wise, Phonetics (Englewood Cliffs, New Jersey: Prentice-Hall, 1957), p. 52.

²²Ibid., p. 49.

²³Peter Ladefoged, "Elements of Acoustic Phonetics," Working Papers in Phonetics (University of California at Los Angeles, 1962), p. 114.

In analyzing a complex wave, called harmonic analysis,²⁴ component wave forms are determined. No sounds for which the acoustic spectrograph is employed are single frequency sounds.²⁵ The sounds are complex ones.

Complex sounds fall into two categories, tones and noise. The distinction between the two is based on their frequency components. These components are sine waves. If composed of a fundamental frequency, the highest common frequency factor of the component frequencies, and a number of other frequencies which are multiplies of this fundamental frequency, overtones or harmonics, the complex sound is designated as a tone. If, however, it is composed of a variety of unrelated frequencies, the complex sound is designated as noise. Speech involves both types of sounds, but the voiceprint technique relies primarily on spectrograms of tones.²⁶

SOUND AND SPEECH

The preceding discussion of sound, resonance, and sound waves plays a significant part in the technique of acoustic spectrography. The relationship of the previous discussion of sound to speech is now examined.

Speech Processes

There are five broad processes involved in the production of speech. First, the brain controls the speech processes. The vocal

²⁴E. Pulgram, Introduction to the Spectrography of Speech (Mouton: Gravenhage, 1959), p. 112.

²⁵Kamine, op. cit., 4.

²⁶Ibid., 4.

cavities, nasal, oral, and pharyngeal, select and suppress overtones; these vocal cavities are the second process. The third is the articulators, lips, teeth, hard palate, soft palate, tongue, mandible, posterior pharyngeal wall, inner edges of the vocal folds, and the hyoid bone.²⁷ The vocal cords are the fourth process. They function to modulate the breath stream. The fifth and last process is the lower respiratory tract, the trachea, lungs, and diaphragm. This process supplies the breath stream. All parts of the speech mechanism interact as a highly coordinated unit.²⁸

Technically, there are no organs of speech.²⁹ There are no organs whose primary function is speech, for speech is an overlaid function. Speech is achieved by means of organs whose primary functions are eating, drinking, and breathing.

Modulation

Speech may be considered as a modulation phenomena.³⁰ Modulation, as used here, means the modification, alteration, variation, or regulation of the breath stream which occurs during speech.³¹ The brain initiates the production of speech by means of certain neuromuscular controls. Under the brain's direction, the breath stream is produced in the lungs. This breath stream is an unidirectional flow of air. A

²⁷H. Kaplan, Anatomy and Physiology of Speech (New York: McGraw-Hill, 1960), p. 272.

²⁸Potter, op. cit., p. 32.

²⁹Wise, op. cit., p. 33.

³⁰"Technical Aspects of Visible Speech," (Murray Hill, New Jersey: Bell Telephone System Monograph, B-1415, November, 1957), p. 5.

³¹Ibid., p. 6.

person speaks by modulating this flow of air which passes through the trachea, vocal cords, and throat, nose, and mouth cavities.

Four types of modulation occur.³² Start-stop modulation is produced by either the vocal cords or articulators. The resulting frequencies are not heard as tones, but as pauses in the flow of speech. The second type of modulation is vocal cord modulation. This method of modulation is produced by the vocal cords which periodically interrupt the flow of air. The result of this modulation is sound waves which give the characteristic of voicing to speech. Vocal cord modulation occurs at fundamental frequencies, 75 Hz to 500 Hz. The fundamental frequency is not only produced, but also a large number of harmonics of the fundamental are generated. The third method of modulation is frictional modulation. Sounds such as s, f, and sh are characteristic of this pattern of modulation. These sounds are known as fricatives. They are produced by the articulators being placed close enough together to form a small opening or constriction through which air must flow. The resulting frictional modulation, caused by turbulent air flow in the constriction, generates a wide range of frequencies that are randomly spaced and harmonically unrelated. The final category is cavity modulation. When the cavities of the throat, mouth, and nose are coupled, they act on the overtones produced by both vocal cord and frictional modulation. These coupled cavities suppress some of the overtones, which, in turn results in an apparent reinforcement of other overtones. In other words, these cavities have the

³²Potter, op. cit., pp. 30-33.

properties of selective transmission and radiation of sound waves.³³

The frequency regions in which the overtones are reinforced are referred to as vocal resonances.³⁴ They are determined by the shapes and sizes of the connected cavities and are changed by movements of the articulators. In speaking, combinations of all the types of modulation mentioned are used.

From Sound to Speech

The modulation typology discussed above is an attempt not only to describe the basis of the production of speech sounds, but also to link the previous discussion of sound to the process of speech.

There are three main points of this linkage. First, the lungs act as a generator. Second, air is again the medium of transmission. Finally, the cavities of the nose, throat, and mouth are the main resonators. For, once the breath stream with its audible frequency components has reached these three cavities, they shape the final product of the speech process. These cavities, depending on the positions taken by the larynx, soft palate, lower jaw, tongue, and lips, selectively transmit and radiate the audible frequency components produced by the vocal cord and frictional modulations.³⁵ Certain of the components or overtones are reinforced, some are damped, and some are completely destroyed. The reinforced frequency regions, vocal resonances, shift in frequency and strength from sound to sound and

³³Ibid., p. 33.

³⁴Ibid., p. 31.

³⁵"Technical Aspects of Visible Speech," op. cit., p. 6.

even during some sounds because of the changes in cavity sizes and shapes which are caused by the movements of the articulators.³⁶

There will be a characteristic mode of vibration of the air corresponding to each position of these vocal organs. As the sounds are formed in succession, the positions of the articulators undergo changes at relatively slow rates and, as a result, mediate the audible overtones in speech. The surfaces of the articulators form the walls of the resonance cavities and change their size and shape.

In an oscillographic analysis of component frequencies of a complex wave pattern, the peaks in the pattern correspond to the basic frequencies of the vibrations. The regions of sound energy around these peaks of relatively large frequency components are called formants.³⁷ The formants are important, for they are directly dependent on the shape of the vocal tract and are characteristic of the particular sound's frequency components.

PHONETICS

Up to the present, the discussion has centered on the acoustic and anatomical aspects of sound and speech. With the introduction of phonetics to the areas of discussion, the last information source is opened for investigation and explanation.

Sounds of a Language: Phonemes and Phones

The configuration of cavities and vocal organs responsible for the variety of speech sounds is dependent on heredity for its individual anatomical contributions. Yet each person can neutralize his

³⁶Ibid., p. 6.

³⁷Ladefoged, "Elements of Acoustic Phonetics," op. cit., p. 92.

own anatomical as well as volitional personal peculiarities to such a degree that each person can learn to make linguistically similar sounds. This similarity is attained when the positions of articulators are generally the same so that similar classes of sounds, called phonemes, are produced.³⁸ A phoneme is the minimum unit of distinctive sound and also a family of sounds.³⁹ A phone is an event of articulation, an articulation which is classified as a phoneme.⁴⁰ For we do not produce phonemes, but, rather, we recognize, classify, and produce each phone within the limits of the classes of phonemes. We learn to observe these limits as relevant to the structured system of any language.

Allophones refer to sounds which are classified as belonging to a particular phoneme, despite differences in the allophone from the ideal phoneme. These differences may result from either the language itself or an individual's idiosyncracies.⁴¹

Types of Sounds

Sounds are phonetically classified by the physiological processes which produce them. In turn, the phonemes of a language are placed in these classes. Various typologies of sound classification are available for use. However, the one presented is considered by the author to be the simplest yet most comprehensive and understandable.

³⁸Pulgram, op. cit., p. 73.

³⁹Wise, op. cit., p. 75.

⁴⁰Pulgram, op. cit., p. 146.

⁴¹James Carrell, Phonetics (New York, New York: McGraw-Hill, 1960), p. 19.

Seven classes are presented.

Vowels. First are vowels. The vowels contain the major sound energy of any syllable. Vowels are tones; their sound waves are regular, repeating themselves in equal time intervals. The wave forms of vowels are complex. The sound itself is heard as a single tone, although it is actually a number of separate elements.

Each vowel sound is distinctive because of its number of frequency components, its frequency, and the relative energy level of its partial vibrations, overtones. The overtones for any given vowel are determined by the manner in which the mouth, throat, and nasal cavities act as resonators. When a speaker produces a vowel, he tunes these cavities in certain ways by muscular adjustments involving articulatory positions. The vocal passage remains open and unobstructed. A person creates different vowels by changing the resonance properties of the cavities through muscular adjustments. Because of their acoustic or sound properties, vowels are known as resonance phenomena.⁴²

Contrary to the normal classification, a, e, i, o, and u are not the only vowels in the English language. The exact number of vowels ranges from between fourteen and seventeen, depending on what sound is included as a vowel.⁴³

Diphthongs. Diphthongs are the second class. These are two vowels spoken together in the same unit of speech. The essential quality of a diphthong depends on the resonance changes. The diphthong is

⁴²Ibid., p. 24.

⁴³"Sound," Collegiate Encyclopedia, ed. Stanley Schlindler (1st ed.; New York, New York: Stratford Press, Inc., 1970), 17, p. 303.

distinguished by the movement from one tone quality to another within the space of the basic speech unit. There are six diphthongs in English.⁴⁴

Semivowels. Semivowels are the third class. The semivowels are sounds which can stand as either vowels or consonants. In some phonetic contexts, semivowels act as vowels; in other contexts, they act as consonants. Their recognition depends on resonance patterns. Nevertheless, the vocal passage is not as open as it is for vowels, since there is always some contact of the articulators which partially obstructs the air flow.⁴⁵

Consonants. Consonants are the fourth class. They are characterized physiologically by some kind of obstruction of the vocal passage or contact of an articulator. Acoustically, they resemble noise. They are traditionally classified as to the degree of closure. If the closure in the throat or mouth is complete, the sound is called a stop. Any stop can be a plosive if the breath, after being compressed, is suddenly released with an explosive puff. The sound of p is a good example. However, if the closure is incomplete, the consonants produced are called fricatives; the sound of f is an example. In the affricatives, a complete closure is followed by a gradual and prolonged release of the compressed breath. The affricative depends on the shift or change for its basic nature and is not a simple stop-plus-fricative combination. The sounds of ch and dg are affricatives.⁴⁶

⁴⁴Potter, op. cit., p. 34.

⁴⁵Carrell, op. cit., p. 25.

⁴⁶Ibid., p. 27.

Glides. Glides, the fifth class are produced by movement of the articulators rather than their static positioning. Acoustically, a glide is a rapid change of resonance which occurs at the phonetic boundary, usually without prominent noise elements. The sounds of w, r, and j, are glides.⁴⁷

Nasals. Nasals are the sixth class. Nasals designate the semi-vowels sounds m, n, ing. The term is chosen because of the distinctive nasal resonance which these sounds contain.⁴⁸

Laterals. Laterals represent the seventh and last class. The sound of l is a lateral sound. The term is used because of the manner in which the voiced breath stream escapes over the sides of the tongue when the sound is made.⁴⁹

Each of the sounds in these classes has its own symbol. The alphabet of these sounds is the International Phonetic Alphabet.

The reason for the preceding discussion of classes of sounds stems from the fact that the sound spectrograph actually breaks down words into the phonemes which make up the words. For example, the word bag would appear on a spectrogram with each of its phonetic sounds readily identifiable. In addition, when the technique of acoustic spectrography is used, what is compared is the basic sound unit, the phoneme. Words, phrases, and sentences are not used. This is the reason why a discussion of the phonetic aspects of acoustic spectrography is necessary.

⁴⁷Ibid., p. 27.

⁴⁸Ibid., p. 28.

⁴⁹Ibid., p. 28.

Moreover, the theory behind the technique of acoustic spectrography, the theory of invariant speech, is also based on the contributions of acoustics, physics, biology, and phonetics.

INVARIANT SPEECH

The theory of invariant speech, as has been stated above, is the basis for voiceprint identification.

The theory of invariant speech is itself relatively simple. The aim of speech is communication. A given group of speakers of a language use a common code and a common set of speech sounds for communication. The speaker tries to produce sounds similar to those produced by others. However, there are differences in the sounds among various speakers. Many submessages that are non-essential to the communication are conveyed along with the essential parts. Vocal anatomy and learned mannerisms of speaking also affect the way in which the sounds are made. In addition, the sounds, submessages, anatomical and environmental aspects are mixed together. The results are twofold. The language and the sounds of the language are similar enough to be understood by other speakers. In addition, spectrographic analysis of the sounds reveals that the frequency positions of the particular phonemes are similar. For example, a sound as the a in the word bag is in a similar position on the spectrogram for all speakers. Phonemes are classes; spectrograms show that these phonemes and the sounds portrayed fall into phonetic classes. What the spectrogram of a phoneme really is is a spectrophoneme. Every phonation of a sound recognizably and typically resembles the same phoneme spectrographically as well as acoustically. This, as may be recalled, was the basis of Potter, Kopp, and Green's work in the visible speech area.

However, the second result is that there are certain differences in the same phoneme among different speakers; and spectrographic analysis shows slight, but definite differences in the frequency position of the phoneme. The theory takes on another aspect when the fact is taken under consideration that there are also slight differences in the sound of the same phoneme when pronounced by the same speaker in succession. No sound is articulated in the exact same way by any number of speakers or by the same speaker saying the sound in succession.⁵⁰

The questions are two: are the variations in the articulations of a sound greater among speakers than the variations within a speaker's articulations; and do the phonetic differences within a speaker's articulations overlap with the phonetic differences among speakers? The theory of invariant speech states that a greater and recognizable difference exists among speakers and that the differences within a speaker are not significant. In addition, there is no overlap between the intraspeaker variability and the interspeaker variability.

Sound Changes

If the problem were only as stated above, then the task of validating the technique of acoustic spectrography would indeed be difficult. It becomes more difficult with the addition of the further complications that not only does the general pronunciation of the sounds of a language change, but also a speaker's articulations may change.

⁵⁰Richard Bolt and others, "On Speaker Identification by Speech Spectrograms: A Scientist's View of Its Reliability for Legal Purposes," A Report to the Technical Committee on Speech Communication of the Acoustical Society of America, 3.

Language Change. The sounds within a language often change with the passage of time.⁵¹ This change may be systematic or unsystematic in nature. The causes may range from physiological limitations on the parts of the human body; to psychological mannerisms influencing the articulation of sounds; and further to the sociological influences of geographic separation.⁵² However, language shifts do not represent such a great problem, because the phonemes themselves remain the same; the words contain different phonemes rather than phonemes having different sounds.

Speaker Articulation Changes. The manner in which a person speaks is determined by both organic and environmental influences. Speech organs only place limitations on speech. Environmental influences, however, do affect pronunciation. A child learns to voice and articulate phonemes correctly by trial and error. The environment influences which phonemes are in which words. Whether the environment influences how phonemes themselves are articulated is a matter of speculation. Again, though, the influence of the environment on what phonemes are in particular words is not a real problem.

The actual problem arises when the articulation of phonemes changes. Old age, for instance, changes the speech organs. There is a loss in the elasticity of muscles used for the speech processes and the cartilages of the larynx do tend to calcify. In addition to the decrease in the usable frequency range, a decrease in resonance in the voice also occurs. Often the voice becomes tremulous because

⁵¹Wise, op. cit., p. 146.

⁵²Ibid., pp. 146-151.

of the reduced integrity of the nervous system.⁵³ New limitations are placed on the articulation of phonemes. Another organic influence is puberty. During puberty, the vocal folds grow. For males, the growth is about 10 millimeters. The lower range of the male child, as a result, drops a whole octave after puberty.⁵⁴ For females, the growth is about 4 millimeters. Her lower range drops two to three notes.⁵⁵ As a rule, voiceprint identification will identify only those who are past puberty. To date no experimental research has been done to determine the actual effects of advancing age. One of the reasons for this is the newness of voiceprint.

Organic changes, then, represent the greatest challenge to voiceprint identification.

Transitions Between Sounds

One final problem confronts the voiceprint technique. This is the problem of the phonetic context. Each phoneme has a distinctively visible pattern on a spectrogram when spoken by itself. However, as these sounds are combined into syllables and words in connected speech, one sound influences and is influenced by the sounds with which it combines. All sounds are not influenced in the same way.

The influences of context result when the articulators change from positions characteristic of one sound to positions characteristic of the succeeding sound. The articulators spend about as much time in the transitional state as they do in the steady-state position, that is,

⁵⁴Willard Zemlin, Speech and Hearing Science (Englewood Cliffs, New Jersey: Prentice-Hall, 1968), pp. 183 and 212.

⁵⁵Ibid., pp. 83 and 212.

in the position producing only a particular phoneme.⁵⁶ The appearance of the resulting spectrographic patterns can be associated with the movement of the articulators from one sound to another.

The basis of how sounds influence each other is the way in which the resonating cavities change shape and size as we say one sound after another in pronouncing words. All sounds are influenced to some extent by the sounds that go before and come after it. The same sounds said together in the same way always look the same, however.

Corollary of the Theory of Invariant Speech.

The technique of acoustic spectrography rests on the theory of invariant speech. Another assumption made is that the sound spectrograph and the spectrograms adequately portray the uniqueness of the individual voice which is the first principle of the theory of invariant speech.

SOUND SPECTROGRAPH

The final section of this chapter deals with the sound spectrograph. Specifically, this section is aimed at illustrating the general, non-technical function of the sound spectrograph, the material the spectrograph is analyzing, and the results the analysis produces. The Voiceprint Laboratories' sound spectrograph is the sound spectrograph described in the following sections.

Function

The function of the sound spectrograph is comparable to that

⁵⁶Potter, op. cit., pp. 38-39.

performed by the middle and inner ear.⁵⁷ These two parts of the human hearing system with its range of 30 to 18,000 Hz analyze the components of a complex sound wave.⁵⁸ The brain identifies the components, thereby identifying the speech sound. The sound spectrograph with its range of 50 to 7,000 Hz takes complex waves, analyzes them, and portrays their component parts.⁵⁹ The sound spectrograph, though, has no distorting prejudices as the human ear has. The spectrogram shows acoustic reality, not subjective perception.

The sound spectrograph is an instrument that analyzes, one band at a time, the simple oscillations of a complex wave and records the amplitude variations in each frequency band side by side in an orderly fashion on a sheet of paper. The result is a visible pattern of sound in its three fundamental dimensions, frequency, amplitude, and time.⁶⁰

The Voiceprint Laboratories' sound spectrograph employs a magnetic tape handling system for continuous erase, record, or playback of a signal from any external source. After a signal has been stored on magnetic tape, the section of tape to be analyzed may be placed on the tape scanner. The scanner reads a 2.4 second section of tape and repeats the signal 400 times in 80 seconds during the analysis. Coupled to the tape scanner is a marking drum and stylus assembly whose function is to record the visual display of the analyzed signal.⁶¹

⁵⁷Kamine, op. cit., 8.

⁵⁸Juris Cederbaums, "Voiceprint Identification: A Scientific and Legal Dilemma," (unpublished paper, New York University School of Law, May 8, 1969), 1.

⁵⁹Operating Manual, Sound Spectrograph Model 4691 A, (Somerville, New Jersey: Voiceprint Laboratories, 1967), p. 13.

⁶⁰Potter, op. cit., p. 11.

⁶¹Operating Manual, op. cit., p. 13.

Basically, the magnetic recording device is used to record the sample of speech to be analyzed. The duration of the sample corresponds to the time required for one revolution of the drum. The speech sample is played back repeatedly in order to analyze its spectral contents. For each revolution, the electronic filter, the scanner, passes only certain frequencies. The energy in the frequency bands activates the electronic stylus so that a straight line of varying darkness is produced across the recording paper. The darkness of the line at any one point indicates how much amplitude energy is present at the specified time within the given frequency band.

As the drum revolves, the passband of the variable electronic filter moves to higher frequencies and the electronic stylus moves parallel to the axis of the drum. Thus, a pattern of closely spaced lines is generated on the paper. This pattern, which is the spectrogram, has the dimensions of frequency, time, and amplitude. On a spectrogram, the frequency is represented on the vertical (y) axis, amplitude by the darkness of the lines, and time duration by the horizontal (x) axis.

The Voiceprint Laboratories' sound spectrograph produces three types of spectrograms, normal or bar, contour, and point-in-time print; but the bar print is the main one of interest for voiceprint identification.

Sound Energy Analyzed

The sound spectrograph analyzes the sound energy recorded on magnetic tape. Of particular importance are the formants. The formants are, again, the areas of concentration of acoustical energy that

especially characterize vowels and semi-vowels. The formants of vowels and semivowels are pictured as dark horizontal resonance bars of definite shapes on a spectrogram. Consonants also have definite patterns.⁶²

The width of the resonance bars is largely determined by differences in the decline in power of the overtones produced by the vocal cavities. The location of the bars is due to the natural frequency resonated by the vocal cavities.⁶³ Differences in the width and location of resonance bars of the same phoneme among different speakers are due to idiosyncracies in every individual's method of speaking. The logical conclusion to be drawn is that idiosyncracies in cavity formation and use cause the differences in the resonance bars. The shape of the cavities is determined by heredity and the placement of the articulators. It is accepted that the anatomical characteristics are unique in each individual.⁶⁴ A controversy arises on the role of the articulators. The theory of invariant speech contends that the positioning of the articulators becomes so habituated that the speaker cannot voluntarily alter his positioning. Hence, the habituation of the positioning of the articulators in conjunction with the hereditarily determined anatomical characteristics produce relatively invariant and unique speech for every individual.

However, as can be seen from persons with speech defects, the speech habits can be altered. The question arises, then, can a person

⁶² Voice Identification Proposal to Office of Law Enforcement Assistance, Prepared by State of Michigan Department of State Police (East Lansing, Michigan, 1968), p. 4.

⁶³ Kamine, op. cit., 16.

⁶⁴ Ibid., 17.

change his habitual speech mannerisms enough to disguise and fool the spectrograph? As yet, no experimental data is available to answer this question. Nevertheless, if, for example, a person had surgery performed to alter his speech mechanism, then there would obviously be a difference in the subsequent spectrogram of his voice. If moreover, a person tried to alter his speech and speech mannerisms by practice, it is a possibility, however small, that he could change the spectrographic portrayal of his voice.

Results of Analysis

The bar spectrogram presents a generally similar pattern for a particular phoneme, irrespective of the speaker who has uttered it. Within the generally similar pattern of a particular phoneme there exists specific personal allophonic differences that characterize a particular speaker. These allophonic differences are consistent for each speaker in spite of the fact that they present a range of variance which hypothetically never overlaps the allophonic range of variance of the same phoneme uttered by another speaker. It is a very remote possibility that two speakers have exactly the same allophonic spectrographic pattern, because of the shapes and sizes and interconnections of cavity resonators and the dynamic use of the articulators. Only a single combination of all these factors among an almost infinite number of possibilities is possible for each phoneme uttered by a single speaker.

A trained person can detect by visual examination the allophonic similarities of spectrograms of a phoneme uttered several times by the same speaker from among spectrograms of the same phoneme uttered by

many speakers. The voiceprint identification technique consists mainly of matching all allophonic similarities in spectrograms of the same phoneme uttered by the same speaker.⁶⁵

CHAPTER SUMMARY

Sound is the basis for voiceprint identification. But it is the sound of the phoneme spoken by a person which is the basic unit for voiceprint analysis. For a person, just as if he had signed his name, imprints his own unique characteristics on the sounds that are classified as phonemes.

Chapter III

A REVIEW OF THE LITERATURE CONCERNING
EXPERIMENTS WITH VOICEPRINT IDENTIFICATION

The statement that a trained identifier can detect by visual examination the allophonic similarities of spectrograms of a phoneme uttered several times by the same speaker from among spectrograms of the same phoneme uttered by many speakers is a one sentence summation of the technique of acoustic spectrography. It is also the central issue of the entire controversy concerning voiceprint identification.

The Michigan State Police have been using voiceprint identification for investigative purposes since 1968.⁶⁶ During the time from September 1968, to November 1969, ninety-eight separate cases were examined by the voiceprint section of the State Police in East Lansing. Forty-six were related to State Police activities and fifty-two were cases brought to the State Police by other state and local police agencies. In all, 120 suspects' voices were tested, and a total of 1,803 spectrograms were examined. Forty-one persons were identified with recorded suspect voices; seventy-nine persons were eliminated or not identified with recorded suspect voices. Some of the identifications were substantiated by admissions of the suspects. The cases ranged from murder, rape, extortion, bribery, and abortion, to burglary,

⁶⁶Ibid., p. 6.

larceny, unauthorized use of an aircraft, threatening phone calls, bomb scares, false reports of crimes, impersonating a police officer, and many obscene phone calls.⁶⁷ Most of the voiceprint cases the State Police have handled have involved telephonic communication.⁶⁸

It must be emphasized that these cases involved the investigative utilization of voiceprint identification - not the identificatory use. The distinction is made because voiceprint identification is in a state of legal limbo. This fact stems from the conflicting experimental data and expert testimony concerning the technique of acoustic spectrography.

This chapter deals specifically with the available literature and research data concerning aural recognition experiments, the experiments of Lawrence Kersta, and experiments of other researchers contesting the validity of voiceprint identification.

AURAL RECOGNITION EXPERIMENTS

The fact that individuals can recognize their friends and acquaintances by their voices is commonly accepted. It has probably happened to everyone at some time or other. Nevertheless, identification of a person by listening to his voice is not always foolproof. Three aural recognition experiments are discussed in the following section.

General Characteristics

Experiments dealing with speaker recognition by listening usually

⁶⁷Ibid., p. 6.

⁶⁸Detective Sargeant Ernest Nash, personal interview, August 10, 1970.

contain three broad characteristics.⁶⁹ First of all, these tests employ the same basic procedure: the speakers are drawn from a prescribed population and are recorded; the recordings are made of the speakers reading selected speech materials; the recordings are edited and presented to the listeners; and the listeners carry out a recognition task. The second characteristic is that each procedural step introduces variables which can influence the resulting performances. These variables include the size and homogeneity of the speaker group, the selection of the speech materials, the size and training of the listener group, the mode of presentation of the recorded speech materials, and the specific task assigned to the listeners. The third characteristic is that the objective of most studies of this type is to test the likelihood that a listener's judgment might be in error.

On Identification of Speakers by Voice

Pollack, Pickett, and Sumby, in 1954, reported their experiment dealing with aural recognition.⁷⁰ Their aim was to see if the accuracy of listeners was aided by the emphasis of some frequencies and the de-emphasis of other frequencies. Pitch was hypothesized to be an important factor in voice identification by ear. Whispered speech was used as an indirect method of testing this hypothesis concerning pitch. In whispered speech, there is very little inflection change.

⁶⁹Summary Review of Procedures for Speaker Recognition, (unpublished paper, Sensory Sciences Research Center, Stanford Research Institute, Menlo Park, California, 1970), 3.

⁷⁰I. Pollack, J. Pickett, and W. Sumby, "On the Identification of Speakers by Voice," Journal of the Acoustical Society of America, 26 (1954), 403-406.

The procedure used was as follows. The researchers chose seven listeners and eight speakers. The speakers' voices were familiar to the listeners because both speakers and listeners worked together in their daily jobs. The speakers' voices were recorded on magnetic tape and played back to the listeners through a high quality audio system. The sound level of the recordings was kept at a relatively constant amplitude. The experimental treatments and the various treatment levels were then presented to the listeners. The task of the listeners was to identify the speaker.

The results of the experiment indicated that:

(1) In aural identification, selection of frequencies to emphasize and de-emphasize frequency levels did not help identification rates.

(2) The more the number of sample voices increased in a trial, the longer the duration of the speech segment to be compared had to be in order to maintain the accuracy rate at its highest level.

(3) Whispered voices also necessitated an increase in the duration of the speech segments of the sample voices. Initially, larger speech segments were needed for whispered speech. As the number of different voices increased, even three times the original duration of the speech segment did not halt the decrease in the accuracy rate. It was concluded that pitch was only a fair influence on aural recognition.

(4) The duration of the speech segment was found to be significant only when the number of voices increased.

Perceptual Basis of Speaker Identity

Voiers reported his experiment on the perceptual aspects of speaker

recognition in 1964.⁷¹ In this experiment, Voiers attempted to determine the number and the nature of basic ways in which voices are perceived by listeners to differ from each other. Voiers was attempting to discover the manners in which listeners perceive the information carried by the voice as useful in identification. He also was attempting to study listener biases or constant errors.

His procedure was similar to Pollack, Pickett, and Sumby's. All the stimulus materials were tape recorded samples of the speech of sixteen men, all born, raised, and educated in the Southwestern United States. Their ages ranged from twenty to thirty-five. All were engineers at one company in Dallas, Texas. The listeners were also natives of the Southwest. They were all college graduates in the area of engineering. Their ages ranged from twenty-three to thirty-five.

The sample of each voice, as heard by the listeners, consisted of a numerically coded identifying phrase, followed by twenty-four sentences of approximately 2.5 seconds duration each, at the rate of one every ten seconds. The listeners were given a semantic differential rating form. This form permitted quantitative representation of a speaker's voice. It consists of forty-nine bipolar items (loud - soft; high - low).

The results indicated that four dimensions contained all the speaker components rated. Clarity, roughness, magnitude, and animation were the four dimensions.

⁷¹W. Voiers, "Perceptual Basis of Speaker Identity," Journal of the Acoustical Society of America, 36 (1964), 1065-73.

Effects of Stimulus Content and Duration
on Talker Identification

Bricker and Pruzansky carried out their experiment at Bell Telephone Laboratories in New Jersey and reported their research in 1966.⁷²

They stated three hypotheses:

(1) Will identification improve with the number of phonemes in a sample of a given duration?

(2) Is the pattern of identifications among talkers independent of the vowel utter?

(3) Do human listeners produce symmetrical misidentification patterns; do particular listeners consistently misidentify particular speakers?

The procedure was similar to that carried out by the other experimenters. Ten males with no accents or abnormalities were recruited from the staff at Bell Telephone Laboratories. Sixteen listeners socially familiar with the speakers were recruited. A soundproof booth was used to make the recordings. A good quality microphone was placed approximately 12 inches from the speaker. The talkers recorded sentences and word lists. The listeners were given a complete day to become familiar with the tape recorded voices of the various speakers.

The results of Bricker and Pruzansky's experiment are shown in Table 1.

⁷²Peter Bricker and Sandra Pruzansky, "Effects of Stimulus Content and Duration on Talker Identification," Journal of the Acoustical Society of America, 40 (1966), 1441-49.

Table 1. Results of the Experiment
of Bricker and Pruzansky, 1966.

<u>Type of Utterance</u>	<u>% Correct</u>	<u>Phonemes</u>	<u>Duration in Milliseconds</u>
Sentences	98	15	2400
Disyllables	87	4	446
Monosyllables	81	3.2	498
Consonant-Vowel Excerpts	63	2	117
Vowel Only Excerpts	56	1	117

As indicated in Table 1, the number of phonemes was more important than the duration of the message, especially with monosyllables. Accuracy of identification improved directly with the increasing number of phonemes sampled from the speaker's repertoire. There may have been an extraneous variable in the differential association of speakers and listeners. The misidentification matrices were very asymmetrical. Some speakers were not missed by all listeners; but each listener missed different speakers. Finally, the vowel uttered did have some affect on the misidentification pattern.

Summary

All of these experiments were far from conclusive or experimentally sterile. Yet, they all tended to point out the fact that the more a listener hears of the speech of a talker, the better the identification accuracy rate becomes. This, again, must be weighed with the fact that all the listeners were socially familiar with all the speakers. In general, aural identification is not foolproof.

SPECTROGRAPHIC ANALYSIS

In this section, the experimental data of Lawrence G. Kersta are presented. These experiments were carried out in his laboratories in Somerville, New Jersey. His claims concerning voiceprint rest on these tests. Problems and defects in the research as cited by other researchers and this writer are stated after the experiment is presented. In the opinion of this writer, most of the experiments described in this section fall short of presenting conclusive data; moreover, many of the experiments fall short of the requirements for good experimentation. Nevertheless, the studies reported here are the main ones pertaining to voiceprint identification.

Kersta's Experiments

Kersta conducted four main experiments to test voiceprint identification. These four experiments are now presented.

First Experiment. Kersta's original experiment, as reported in Nature,⁷³ consisted of matching tasks with spectrograms. Matching tasks employ a sorting test; that is, there is always a match in a particular trial. There is a certain number of known speaker samples. A spectrogram of a supposedly unknown suspect can always be matched with one of the known spectrograms. The task is to pick out the right known spectrogram that matches with the unknown spectrogram.⁷⁴ No true identification is made. A discrimination task is a better description

⁷³Kersta, "Voiceprint Identification," op. cit., 1253-57.

⁷⁴Bolt, op. cit., p. 9.

of matching. For, the question of is this or is this not the man, is never asked. Rather, the question of which one, among so many, is the same, is asked. The identifiers know that there is always a match in any particular trial.

Kersta asked the question, how well could an unknown speaker be identified with the correct, known speakers in a definite population of speakers? He chose eight high school girls as his identifiers. They were sixteen to seventeen years old. Each was given a week of training in voiceprint reading and detection of unique clues in voiceprints. The girls worked in panels of two.

The procedure was to present four utterances of one word spoken by 5, 9, and finally 12 speakers. Ten common words in all were used for the utterances - it, is, on, you, and, the, I, to, me, a. Only one word was used at a time. Four spectrograms from each speaker used were then presented to the girls in a random order. As each was given, they were to sort it into a pile for each speaker. The girls were told how many speakers there would be, 5, or 9, or 12, and each speaker in a group had an equal number of spectrograms. For example, when 5 speakers were used, 20 spectrograms were presented; with 9 speakers, 36 spectrograms; with 12 speakers, 48 spectrograms. In sorting trials with 9 and 12 speakers, over 2000 matching tests were run.

The results of Kersta's experiment are illustrated in Table 2.

Table 2. Results of Kersta's Experiment, 1962.

<u>Type of Print</u>	<u>Words in Isolation Range of Error in %</u>	<u>Range of Error in % for Single Words</u>
Contour	.37 to 1.5	
Bar	.35 to 1.0	
Bar and Contour		0.0 to 1.8
Total Error = .8%		

For words taken from context, the total error was 1.0%.

Kersta used this experiment to unveil his new technique of voice identification in the Nature article. The foremost problem with the experiment was that it was not an identification experiment; it was a matching test. Secondly, the identifiers knew the number of speakers and that each speaker had an equal number of spectrograms. Third, there were too few speakers. A practical task may involve a search of 100 to 200 spectrograms.⁷⁵ Finally, the exact procedures used in this experiment were never specified in the Nature article or in any other publication.

Kersta kept within his hypothesis in carrying out the experiment. His generalizations from the results, however, were premature. As an experiment, it was useful. As proof of the workability of this technique, it was only the first step.

Second Experiment. In the second experiment,⁷⁶ Kersta used two lists of five words:

⁷⁵Mr. Ralph Turner and Mr. Clarence Romig, personal discussion, September 28, 1970.

⁷⁶Lawrence G. Kersta, "Voiceprint Infallibility," (unpublished paper, presented to the Acoustical Society of America, November 7, 1962), 1-3.

1. To, me, and, the, that.

2. A, I, is, on, you.

He used the steady-state portions of these words taken from context.

Again, the high school girls were used.

The presentation procedures were not exactly the same as in the first experiment. The girls were given five spectrograms of a single speaker showing the five words from one list. Their task was to find the match from a group of fifty piles of spectrograms. Each pile contained five spectrograms of the words from the list being used at any particular time.⁷⁷

The results showed that for list 1, the mean range of error was .94%; for list 2, it was .75%.

As in the first experiment, the exact experimental procedures were not enumerated. The test used was more like an identification task since the girls had to choose one sample out of fifty; but the procedure was still more a discrimination than identification task. The match was always in the fifty samples.

Third Experiment. Five fingerprint experts were used after one week of training to replicate the second experiment. Their accuracy rate was 93.4%.⁷⁸

Fourth Experiment. Kersta next attempted to determine if either fraternal or identical twins were not subject to the theory of invariant speech.⁷⁹

⁷⁷Kamine, op. cit., 19.

⁷⁸Ibid., 19.

⁷⁹Lawrence G. Kersta and J. Colangelo, "The Spectrographic Speech Patterns of Identical Twins," (unpublished paper, Voiceprint Laboratories, Somerville, New Jersey), 1-4.

Fifteen pairs of male and fifteen pairs of female fraternal twins were recorded and their voices were used to train two previously untrained high school girls. Then recordings of the sentences, you were, and, were you, were made by thirty pairs of identical twins for the actual experiment. The pairs of identical twins ranged in age from seven to twelve years old.

A matching procedure was used to test the identifiers. One pair of twins was tested at a time. Each identifier had one sample spectrogram of each of the twins' utterances. The identifiers were then given a single spectrogram of the sentence, you were, and told that one of the two twins whose samples they had in front of them had uttered the sentence. The identifiers were to match the correct known sample to the unknown sample.

For identical female twins, the accuracy rate was 84%; for twin males, it was 90%.

The previous criticisms of this type of experiment also apply here. What Kersta was attempting was to show that identical twins do not have identical voices and spectrographic patterns. Considering that only two words were used to match the known and the unknown, the results could be considered to be significant.

Miscellaneous Tests

Kersta has done other experiments along the same lines as those presented above. He has enlarged his voice file to 123 speakers.⁸⁰ However, his experimental techniques have remained the same.

⁸⁰Lawrence G. Kersta, "Voiceprint Identification," Newsletter of the American Academy of Forensic Sciences (June-July, 1970), 25.

One more of Kersta's experiments is presented. This test was one he devised for use with automated identification.⁸¹ The technical method of coding he devised is not presented, for it is not necessary to understand it to understand his study.

Basically, he devised a numerical, computer-like code for contour spectrograms. The voice samples of forty-six speakers were coded. 1,012 trials were run for each speaker. Kersta's results indicated that there was a 2.2% error incidence for an utterance made at a later date not to meet the code made for the original voice sample of the same voice and the same word. In addition, there was 4.35% chance of duplication; that is, there was a 4.35% chance that the same code could result for two different people. Finally, there was an 11.9% mismatch error rate. However, only one word for each speaker was used for this coding. If five words were used, the mismatch error rate would be reduced to 2.3 per 100,000.⁸²

This experiment, for all its technical aspects, was much more sophisticated than any of his others. This technique would seem to offer some promise. Computerization of his coding procedure could open the way for a realistic model of computer identification by spectrographic analysis.

Kersta had made other tests of voiceprint identification which were not of an experimental nature. He has tested Sherry Lewis,⁸³

⁸¹Lawrence G. Kersta, "Voiceprint Classification," (unpublished paper presented to the Acoustical Society of America, February 14, 1967), 1-6.

⁸²Ibid., 5.

⁸³Asp, op. cit., 3.

a ventriloquist, as well as Elliot Reed and Vaughn Meder,⁸⁴ both mimics, to show that no matter how skilled a person is, no one can imitate another person's voice so well that the spectrographic patterns are the same.

Section Summary

As can be seen from the reporting of the above experimentation and tests, a comprehensive and tightly controlled experimental test of the technique of acoustic spectrography is yet to be done. Kersta used results from his experiments to propose that voiceprint is a valid technique for positive identification of individuals by their voice. No one challenges the assumption that the sound spectrograph analyzes a person's voice. However, that positive identifications of suspects by their voices are possible has been and is challenged by phonetic experts.⁸⁵ Kersta's experiments have also been attacked as not being relevant to positive identification. His manner of publication, not specifically outlining his procedures, has also come under fire. Finally, the sophistication of the sound spectrograph to provide enough reliable information to permit positive identifications even if the theory of invariant speech were proven has been challenged.⁸⁶

CONFLICTING DATA

Peter Ladefoged, a phonetics expert, has made these statements:

⁸⁴Bolt, op. cit., 2-3.

⁸⁵Peter Ladefoged and Ralph Vanderslice, "Voiceprint Mystique," Working Papers in Phonetics (University of California at Los Angeles,

⁸⁶Kamine, op. cit., 26.

"The same speaker saying the same word on different occasions without an attempt to disguise his voice may produce spectrograms which differ greatly...."⁸⁷

"The hypothesis which we would like to propose at this point is that a speaker may pronounce a given word in different ways on different occasions; and that the range of one speaker's pronunciations may overlap with those of another so that different speakers may be responsible for spectrograms which are very similar."⁸⁸

Two major experiments are presented whose results purport to show that voiceprint identification is neither infallible nor better than aural recognition. The results of these experiments cannot be cast aside; yet, these tests, run by reputable and well known researchers, are not infallible and conclusive in themselves. An evaluation of each experiment is presented as was done after each of Kersta's experiments.

Effects of Context on Talker Identification

Young and Campbell are researchers who attempted to retest Kersta's method of voiceprint identification.⁸⁹ The identifiers for this experiment were seven Ph.D. candidates in Speech Pathology and Audiology; three were assistant professors in the same field. All were familiar with spectrographic analysis. The basic strategy was to train the identifiers to be able to identify speakers on the basis of words spoken in isolation and then to have them identify the same speakers using the same words, but spoken in context. The hypothesis was that

⁸⁷Asp, op. cit., 6.

⁸⁸Ladefoged, "Voiceprint Mystique," op. cit., p. 131.

⁸⁹Young, op. cit., 1250-54.

if each person's voice was truly unique, then the accuracy level of identification performance for words in different contexts should be similar to the accuracy for identifying talkers from words spoken in isolation.

Each identifier was individually trained and tested by either Young or Campbell. Initially, the researcher gave the identifier a general description of what was to come. Spectrograms of four repetitions of the word, me, spoken by five different speakers were presented to the identifier. In this way, the similarities and differences could easily be seen. With the twenty spectrograms in front of the identifier, the researcher listed, described, and discussed the acoustic cues thought to be most associated with identifying speakers by the voiceprint technique. Those described included:

- (1) Frequency, intensity, and bandwidth of the speakers' spectrograms.
- (2) Transitional characteristics - the relatively rapid change in frequencies that occurs between certain phonemes.
- (3) Asynchronous onset and termination of the first three resonance bars of the formants.
- (4) The frequency spacing between the formants.
- (5) The presence and characteristics of interformant resonances.
- (6) The purity and regularity of the vertical striations as a degree of periodicity of the voice.
- (7) The horizontal distance between the vertical striations as an indication of the fundamental frequency.

(8) The overall duration of the word production.⁹⁰

The total training time for an identifier ranged from one and three-quarters hours to two and one-half hours.

The words used were the following: it, is, the, you, a, I, and, to, me, on. The sentences from which the key words were taken from context were:

But you said it first. Do you see it now?

Is it really you? Yes, it fits you.⁹¹

The underlined words were the words taken from context. There were eight sentences in all, but only the four with the underlined words were used. The speakers were five normal, male adults.

Spectrographic analysis was done by a Kay Sonagraph, model 661 A. Four different recordings and sonagrams of each word and sentence were made. All the samples were recorded and reproduced at the same levels.

The first part of the experimental tasks involved a matching test. Each word was taken individually. One of the four repetitions from each of the five speakers was given to the identifier. Then the other three repetitions from each of the five speakers were randomly given to the identifier to match with the samples. This task was done four times for each word. These were considered to be the training tasks.

Immediately upon completion of these training tasks, the actual experimental tasks were presented. Again, these were matching-to-

⁹⁰Ibid., 1251.

⁹¹Ibid., 1251.

sample tasks. The experimental tasks were different from the training tasks in that each known sample showed the words, you and it, that had been spoken in one of the four sentences containing these words. The samples to be matched were the remaining three sets of repetitions of these sentences.

The identifiers were encouraged to take as much time as they needed to make a correct match. They were paid fifty cents immediately after each correct match.

The results of the experiment of Campbell and Young are listed in Table 3.

Table 3. Results of Campbell and Young's Experiment, 1967.

<u>Observer</u>	<u>Words in Isolation</u>	<u>Words in Context</u>
	Mean Scores in %	Mean Scores in %
L.M.	80.0	53.3
D.S.	77.3	53.3
J.L.	76.7	46.7
L.B.	78.9	46.7
W.G.	85.6	40.0
R.I.	73.3	33.3
J.N.	86.7	33.3
A.C.	74.4	26.7
M.P.	78.9	20.0
D.M.	75.6	20.0
Average	78.4	37.3

In the discussion of the results, Campbell and Young list four main conclusions:

(1) With words spoken in isolation, the observers passed the 0.0001 criteria level. For words in context, the best done in one trial was eight out of fifteen correct; this verges on chance occurrence.

(2) Words spoken in context were of shorter duration than the same words spoken in isolation. Less acoustic information was transmitted.

(3) The phonetic environment of the test words interacted with the words' acoustic representation; overlapping physiological sequences occurring in running speech result in overlapping and interacting acoustic parameters.

(4) The last two effects, word duration and phonetic interaction, seemed to have outweighed any intra-talker consistency.

The experimenters also discussed why their results differed from Kersta's research. Second, the cues used were those which seemed to Cambell and Young to be the ones most likely to have been used by Kersta's trainees. However, they admitted that they did not know exactly what cues Kersta did use. Identifier characteristics may have had some affect; this was the third possible reason for the discrepancy. Finally, their method of experimentation, while it was a matching task as Kersta's, may not have given as many clues as did Kersta's.

There are five additional reasons why the results differed so much from Kersta's. One, the basic hypothesis was theoretically incorrect. As was explained earlier, the phonemes spoken in isolation have a definite and characteristic appearance. But connected speech does influence the phonetic appearance on a spectrogram. Unless this is taken into consideration, the task of identifying sounds in context becomes completely different from identifying sounds in isolation. If the influences are considered, then the in-context tasks become an

extension of the in-isolation tasks. Two, the training time was far too short. Three, the cues used were not all correct. For example, the duration of the sound has very little bearing in the technique of voiceprint identification. If an identifier based his judgment on this cue, his conclusion would be based on erroneous information. Four, it was not reported how the words in context were marked or cut out of context, or if this was done at all. The marking or cutting is done to indicate when a sound is starting and stopping. Not only does this procedure take much experience to perform correctly, but also if it is done inaccurately, the transitional influences may be used rather than the steady portions of the phoneme: the identifier may be trying to match two different sounds. Five, the training experience may have actually hindered rather than have helped the identifiers. They were trained for one type of word, and tested on another type.

Speaker Authentication and Identification

Stevens, Williams, Carbonell, and Woods carried out the second major experiment.⁹² These experimenters added a new feature to their tests. They used the matching-to-sample tasks, but varied from the model. Instead of the unknown speaker always being in the known pile, they added speakers to the unknown pile who were not in the known pile.

These researchers recruited twenty-four speakers, 20 to 40 years old. None had any accent. Six identifiers were also recruited. These were three men and three women, all college students.

⁹²Stevens, op. cit., 1296-1607.

Nine words, one phrase, and one sentence were used. The nine words were the following: baseball, sidewalk, pancake, dovetail, yardstick, scarecrow, that, base, side. The phrase was: a baseball glove. The sentence was: That sidewalk is broken.

The equipment used to record and spectrographically analyze the speakers by these researchers were pieces of equipment combined to construct their own spectrograph.

The tests consisted of aural and visual identification. The first part of the tests was strictly matching-to-sample type tasks similar to Young and Cambell's. The second half was different, for unmatchable speakers were added. The identifiers had to first determine whether the sample to be matched was in the known category at all. If they decided that it was in the known category, they then tried to match it. If they concluded that it was an unmatchable sample, they discarded it.

The results of the tests were that aural recognition proved to be superior to visual recognition. For normal match-to-sample tasks, aural had an 18% error rate; visual had 28%. For the second type of match-to-sample, the error rate for aural recognition was 6%; the rate for visual was 21%.

The researchers concluded that, first, aural recognition is superior to visual recognition. Second, they concluded that performance improves with experience more rapidly for aural than for visual recognition. Finally, they concluded that longer utterances increase the probability of a correct visual identification.

An evaluation of this experiment results in the conclusion that if any training in visual identification was given it was not reported. In addition, the equipment used may also have influenced the visual samples; these researchers used neither Voiceprint Laboratories nor Kay Electric's sound spectrograph. Because they did not use the standard sound spectrographs, whatever information the spectrograms contained or did not contain is in doubt. No identification cues for visual identification were stated. In general, it appears that the identifiers were merely handed ersatz spectrograms and instructed to compare them and attempt to make an identification.

CHAPTER SUMMARY

The most obvious conclusions to be drawn from the above cited experiments are that the experimentation reported has not been of the highest caliber; and the results are not only conflicting, but a consequence of the defects in the experiments.

Aural recognition is not completely dependable. Kersta's experimental data tends to show that the technique of voiceprint identification might be reliable. The experiments of Young and Cambell, and Stevens, Williams, Carbonell, and Woods propose to prove that aural recognition may be more reliable than voiceprint identification.

The Acoustical Society of America took this stand in 1966:

"The technical committee on speech communication is concerned that voiceprints have been admitted as real evidence on the basis of claims which have not yet been evaluated scientifically. The committee invites the executive counsel to consider the matter and take appropriate action."⁹³

⁹³Asp, op. cit., 5.

As far as the scientific community is concerned, this is the official status of voiceprint identification.

Chapter IV
RESEARCH OF THE SCHOOL OF CRIMINAL
JUSTICE OF MICHIGAN STATE UNIVERSITY

In 1967, the Department of Justice of the Federal Government funded a grant requested by the Michigan State Police to carry out research on voiceprint identification.⁹⁴ As a subcontractor to the State Police, the School of Criminal Justice of Michigan State University was commissioned to assume part of the research.

This chapter contains the description of this research, carried out in 1969 and 1970, a statement of the results, a discussion of those results, and a brief comparison of the initial results of the research project of the Audiology and Speech Sciences Department of Michigan State University.

THE RESEARCH PROJECTS

Goals

The goals of the research of the School of Criminal Justice were six:

(1) "Assemble pertinent data regarding Voiceprint Identification from the Department of Audiology and Speech Sciences, Michigan State University, and the Michigan State Police."⁹⁵

⁹⁴Voice Identification Proposal, op. cit.

⁹⁵Ibid., p. 5a.

(2) "Develop practical field test situations in cooperation with the Department of Audiology and Speech Sciences, Michigan State University, and Michigan State Police."⁹⁶

(3) "Conduct field tests, simulating practical situations wherever possible."⁹⁷

(4) "Submit field test evidence for personal identification, again simulating practical circumstances wherever possible."⁹⁸

(5) "Draw conclusions from field tests."⁹⁹

(6) "Considering the aforementioned goals, it is accepted that the major objectives of the field evaluation are to lay a foundation for further studies."¹⁰⁰

The hypotheses to be tested were two. (1) It was hypothesized that the correct identification of one speaker from many speakers could be made in a sufficient number of tests from a sufficiently large sample population. (2) If the technique showed promise of being valid and reliable, it could be applied not only in laboratory research but also in practical situations.¹⁰¹

Planning

Two specific goals were established. First, actual field tests would be undertaken to gather data. These tests were to be in situations as practical as possible. Banks, stores, gas stations,

⁹⁶Ibid., p. 5a.

⁹⁷Ibid., p. 5a.

⁹⁸Ibid., p. 5a.

⁹⁹Ibid., p. 5a.

¹⁰⁰Ibid., p. 5a.

¹⁰¹Ibid., pp. 5b-5c.

dormitories, and office buildings were possibilities. Second, the training of two undergraduates from the School of Criminal Justice in the technique of voiceprint identification was decided to be initiated. This course of action was decided upon for three reasons. One, the experiments with the gathered voice samples and spectrograms required that identifiers be available for the specific identification tasks to be carried out by the School of Criminal Justice. An extensive period of time for these identifications would be necessary. Identifiers working in the Audiology and Speech Sciences' voiceprint research project were available for use by the School of Criminal Justice, but the amount of time required for the identifications to be run would have interfered with the Audiology Department's research. Two, assistance in gathering the voice samples and aid in other areas were needed. If identifiers were to assist in this work, valuable experience and increased understanding of the process and theoretical background of the voiceprint process would be gained by them. Three, by training two additional identifiers, the actual experiment in identification could be designed to be a part of the total design of the research. Moreover, proper control of the identification procedures could be maintained if the identifications were done under the control of the same research project which gathered the voice samples, made the spectrographic analyses of the samples, and marked and coded them. Continuity in the research process could be maintained. One additional benefit of training two separate identifiers and running separate identification experiments was that a comparison of the School of Criminal Justice's project results with the Department of Audiology and Speech Sciences's project results would be possible. The training

of the identifiers of the School of Criminal Justice would be provided by the director of the voiceprint project of the Audiology and Speech Sciences Department. The director had received two weeks of training from Lawrence G. Kersta in Somerville, New Jersey.

The design of the experiment was divided into three sections. The training of the identifiers was to be completed prior to the gathering of the voice samples. Following this a pilot project was to be carried out in order to gain experience for the main experiment and to determine what practical and technical problems there would be. The information gained from the pilot study would then shape the final section of the research, the actual experiment.

Identifiers

Two identifiers were chosen. One was a senior in the School of Criminal Justice and a criminalist major; the other was a general law enforcement major. The first identifier chosen had already received training and had worked in the Audiology and Speech Sciences Department's voiceprint project. The second had received no training. He had seven years experience in law enforcement.

Training

The total training time for the one identifier was six weeks. The first three weeks were used for lectures. These lectures were given by the director of the Audiology Department's voiceprint project. The subjects of the lectures were phonetics, the history of language, the theory of invariant speech, the sound spectrograph, the operation of the sound spectrograph, and acoustic spectrography. Six hours a week were devoted to lectures.

The last three weeks were devoted to practical tasks. The director of the voiceprint project of the Audiology Department also gave the instructions concerning the application of the technique of acoustic spectrography. These tasks were true identification tasks. One spectrographic exemplar was given to the identifier and as many as 40 spectrographic examples of other voices were presented to him. He was to determine whether or not there was a match to the exemplar in the 40 samples. The matching voice was not necessarily among the 40 samples. The problem was to identify the one speaker. If no match was found, then no match was the answer. An absolute judgment was called for.

The practical identification tasks were started on the simplest level, words in isolation. The difficulty of the spectrograms was increased step by step. From words in isolation, the tasks progressed to words in a fixed context. Fixed context means that the words were spoken in the same order in all the spectrograms. For example, if the phrase, "the money in the bag," was the sample phrase, the words in the phrase would always be in the same order for all speakers. The next step was words in random context. In this type, the order of the words was not fixed; rather they were in a random order. For example, "the money in the bag," might be spoken as "bag, money, the, in." All of the steps, including the last, involved contemporaneous speech recordings; that is, the sample spectrograms and the exemplar spectrogram were made of recordings taken within a day interval of each other. The next step in the degree of difficulty involved non-contemporaneous spectrograms. Non-contemporaneous spectrograms were spectrograms of speech samples in which the sample to be matched was more recent than

the comparison samples. These are more difficult than contemporaneous samples because of the intra-speaker variations. The intra-speaker variations are more noticeable when the speech samples are taken, for instance, two months apart than when they are taken two hours apart. Using non-contemporaneous spectrograms, the words in isolation, words in fixed context, and words in random context steps were repeated. The spectrographic voice exemplars which were used for training were made of direct speech recorded by a tape recorder.

There were eight cues on a spectrogram used to detect a matching spectrogram:

- (1) The patterns of the frequency components, especially the amplitudes (the stress put on a particular phoneme, for example).
- (2) The frequency band widths of the resonance patterns.
- (3) The overall shapes of the resonance bars or formant patterns.
- (4) The slopes of the resonance bars.
- (5) The mean frequencies of the resonance bars.
- (6) The separation of the resonance bars.
- (7) The interformant spectra (the resonance patterns between formants).

- (8) The vertical striation patterns.¹⁰²

These tasks were done for two to three hours a day, four days a week for three weeks.

¹⁰²Oscar Tosi, "Speaker Identification Through Acoustic Spectrography," (unpublished paper presented to the XIV Congres De Logopedie et Phoniatrie), 3.

At the end of this period of six weeks, the identifier had reached a level of proficiency which was acceptable to the director of the Audiology Department's voiceprint project. It was at this time that the next phase of the design was implemented, the pilot study.

Pilot Study

A practical field situation was chosen. However, certain methodological questions arose; for examples, how to record the sample voices and where to record these voices had to be decided. Banks, grocery stores, gas stations, department stores, drug stores, and college dormitories were all possibilities. It was concluded that a dormitory was the best site for the first recordings of the unknown (field) samples. The pilot study was undertaken in the main lobby of the dormitory. This area was chosen because it best suited the criteria of the research design; it had all the characteristics of a public building. First, the area used had a large number of people in it; this was so that a suitable sample of individuals could be drawn from for the voice samples. The individuals in the dormitory were all over twenty-one years of age. Second, the cooperation of the people in the dormitory could easily be obtained. Third, the cooperation of the manager of the dormitory was granted for the research project. Finally, the atmosphere of the lobby was like that of any office of a business or bank. There was the background noise of a cash register, typewriters, people walking by, and people talking.

Two Uher portable tape recorders¹⁰³ were obtained from the State Police. These tape recorders were the ones actually used by the State Police to record for their cases.

Procedures. The procedures for the voice gathering at the dormitory were the following. A table was set up in the lobby of the dormitory. One Uher recorder was placed on this table. The other recorder was taken to the reading room in the basement of the dorm. The reading room was chosen as the quiet room to record the laboratory (known) samples. Because it was rarely used by anyone during the daytime hours, and because there was very little noise in the room, it was ideal for recording. One technician was stationed at the table in the lobby. His role was to operate the tape recorder and record on a card the name and speaker number of the speaker, as well as such information as nationality, and age. Another assistant was to ask individuals as they walked through the lobby to volunteer for the research project; he also was to operate the tape recorder in the quiet room.

If a person volunteered to be recorded, then he would be taken to the table in the lobby at which a technician was stationed. The volunteer would then be given a brief explanation of what the project was about and what part he was to play in it. If he had any questions, then they would be answered. The technician would ask for the individual's name, room number, nationality, and age. The sex of the person was also recorded. Next, he would be told what he was to do. Generally, this was simply to state their speaker number and read the two sentences twice while the tape recorder recorded his speech. After

¹⁰³See Appendices B and C for Specifications.

this part was completed, the speaker was asked to accompany the assistant down to the quiet room. In the quiet room, the same samples of the speech of the speaker were again recorded. The same process was repeated for all speakers.

Speech Samples and Speakers. The four sentences which were used were:

This is a hold-up; put the money in the bag.

This is a stick-up; give me the money.

Give me the money; and put it in the bag.

Put it in the bag; and take it with you.

The first two sentences were used for only the first two volunteers because some anxiety was aroused by their usage. These two sentences were discontinued.

The volunteers were 12 males and 8 females. Originally, the plan called for an equal number of males and females. However, women were found to be less cooperative and more suspicious than males. The fact that all persons were told that the research was under the direction of the School of Criminal Justice was one reason why some individuals became suspicious of our motives.

The recordings took two days to complete. This was much longer than expected because of the time at which the recordings were made, a vacation period. Very few American students were present. Most of those students recorded were from Asia, particularly Thailand and Nationalist China.

The three difficulties originally encountered, then, were the type of sentence, the lack of enough students, and the suspicion raised by the sponsorship of the School of Criminal Justice.

Recording Tapes and Spectrogram. All the reels of tape, standard magnetic tape, were marked with the date of the recording, the place, and the number of speakers who were recorded on the particular tape.

The tapes were then taken to the Audiology and Speech Sciences Department. There spectrograms were made of all the speech samples. The spectrograms were then marked. The words spoken were written in underneath their spectrographic representation. These words were in a fixed context. The spectrograms were then coded, using the Rand Corporation's book of 1,000,000 random digits.¹⁰⁴

A week after the first recordings were made, six of the original speakers were contacted, and appointments were made for them to re-record their speech. The same procedure as before was followed for these recordings and spectrographic analysis.

Identification Tasks. The identifiers were then instructed on how, where, and when the identification tasks were to be run. The voiceprint office of the School of Criminal Justice would be used. Specific times during the week were specified. The identifiers were to work separately.

The identification tasks were real identification tasks, not matching-to-sample tasks or discrimination tasks. An absolute judgment was demanded. Either a match was present or it was not.

The identification tasks were set up as follows. There were thirty tasks. First, each task was planned to contain 9 unknown speakers and 5 known speakers. The identifier was handed the 9 unknowns and then

¹⁰⁴Rand Corporation, A Million Random Digits with 100,000 Normal Deviates (Glencoe, Illinois: Free Press Publishers, 1955).

handed the 5 knowns. A speaker's spectrograms were both repetitions of the two sentences. The average number of spectrograms of a speaker handed to an identifier was four; sometimes more than four were given. However, no pattern was made by the number of spectrograms of a speaker given to an identifier. The identifier was told that there could be males and females in the speaker samples, matches and no-matches, and double matches. The double matches meant that in the known spectrograms given to an identifier there could be the voice of the same speaker twice. There actually were double matches given. The non-contemporaneous recordings were used for this purpose. The identifier was then instructed to take one of the five known speakers at a time and not to rule out a speaker in the unknown group because he had been identified once with a known speaker; the possibility of a double match was present. The possibility of double matches was put into the tasks to serve two purposes. First, if double matches were possible, no speaker in the unknown group could be ruled out because he had been identified once from a group of five knowns; the same voice might be in the same group of knowns a second time. Secondly, because the number of speakers was too small to permit more than 9 unknowns to be presented at one time, the double matches made fewer spectrograms necessary. If no double matches were put in, then a new group of unknown spectrograms would have to be presented each time a new known was presented. Otherwise, the identification tasks would degenerate into a process of elimination. With double matches, the same 9 unknowns could be used for the 5 knowns presented in one trial.

Results. The results of the pilot study are illustrated in Table 4.

Table 4. Results of the Pilot Study of the School of Criminal Justice. Percentages and Raw Scores are Shown.

<u>Identifier</u>	<u>Matches</u>	<u>No-Matches</u>	<u>Total</u>
First Identifier	66.7% (16/24)	83.3% (5/6)	70% (21/30)
Second Identifier	62.5% (15/24)	100.0% (6/6)	70% (21/30)
Total	64.6% (31/48)	92.5% (11/12)	70% (42/60)

A Chi square statistical analysis was run on the data. Table 5 illustrates the results of this statistical analysis.

Table 5. Results of the Chi Square Statistical Analysis. Results Shown Indicate Level of Significance.

<u>Identifier</u>	<u>Matches</u>	<u>No-Matches</u>	<u>Total</u>
First Identifier	.01	.01	.01
Second Identifier	.01	.01	.01
Total	.01	.01	.01

The results obtained from the pilot study indicated that, while the level of significance reached by the identifiers for correct number of identifications was statistically significant, the 30.0% chance of error presented a real problem. One of the factors which may have been responsible for this low rate was the considerable amount of background noise that appeared on the spectrograms. Both identifiers remarked that the resonance spectra that appeared on the spectrograms made it difficult to distinguish the voice resonance from the noise resonance.

The main conclusion drawn from the pilot study concerned the amount of background noise. It was believed that the background noise would adversely affect the accuracy rate unless the amount of noise that was recorded could be reduced. It was extremely difficult for the identifiers to use the visual cues and information normally present on a spectrogram when noise patterns obscured and gave false information to the identifiers.

The design and procedures of the main study benefitted from the pilot study. Three points have already been mentioned, the problem of enough speakers, the suspicions of some speakers, and the type of sentence to be recorded. The effects of the background noise were also valuable knowledge.

PRINCIPAL STUDY

The Cashier's Office in the Administration Building of Michigan State University was selected as the site for voice gathering. It is a functioning bank; there were large numbers of students making daily use of the facilities; and the cooperation of the administrators was more likely than at any other local bank.

Factors Under Study

For the main study, three factors were under evaluation. The sex of the speaker was the first factor. Although some writers have stated that no problems have been encountered with spectrographic analysis of

females voices,¹⁰⁵ Malmberg¹⁰⁶ and Pulgram¹⁰⁷ have stated that female voices, because their shorter vocal folds vibrate faster than male voices, have about 17% higher formant frequencies than males.¹⁰⁸ Because of the discrepancy in these writers' opinions, the sex of the speaker was chosen as one factor for experimentation.

The second factor was the type of recorder used. The two Uher recorders borrowed from the Michigan State Police were used again for the main study. Two Wollensak portable tape recorders, model T-2000, were also used. The factor here was the type of recorder. Only these two types of recorders were used. The study attempted to determine if either make of recorder was better than the other.

The third factor was the type of microphone used. The Uher recorders came with their own microphones, model 515. Two different microphones were chosen for the Wollensaks. An Electro-voice dynamic, omnidirectional microphone, model 6358, was chosen for one Wollensak; an Electro-voice dynamic, unidirectional microphone, model RE 15, was chosen for the other Wollensak. The assumption here was that the Wollensak with the dynamic, omnidirectional microphone would be equivalent to the Uher recorder and microphone. If any difference in accuracy levels of identifiers was found between the two recorders, it

¹⁰⁵Thomas Coon, "Voiceprint Identification," Police, 10 (May-June, 1966), 91.

¹⁰⁶Pulgram, op. cit., pp. 131 and 145-146.

¹⁰⁷Bertil Malmberg, Manual of Phonetics (Amsterdam: North-Holland Publishing Company, 1968), pp. 181 and 233.

¹⁰⁸Pulgram, op. cit., p. 131.

could be attributed to the recorder model.¹⁰⁹ If no difference resulted, then the two recorders and their microphones could be considered as equal. The next comparison would be between the dynamic omnidirectional microphone and the dynamic, unidirectional microphone. It was because of the large amount of background noise that the unidirectional microphone was picked. It was hypothesized that the qualities of the microphone would eliminate some of the background noise and result in clearer spectrograms, thus increasing the accuracy level of the identifiers.

Procedures

A teller's window in the Cashier's Office served as the site of the field recordings. Three tape recorders, the two Wollensaks and one Uher, were placed in the teller's booth. The microphones were placed on a cardboard stand in the middle of the top of the counter area, facing directly outward at an upward angle of 45 degrees. The positions of the microphones were randomly changed four times to prevent any one microphone from being on one side or in the exact middle all the time. The recorders were run at a speed of 7 1/2 inches per seconds. The recording levels were kept at a constant level.

The other Uher was taken to a stand in an alcove of the Cashier's Office. This Uher recorded the laboratory type samples of the known spectrograms.

There was a large amount of noise from typewriters, change machines, adding machines, and people talking, entering, and leaving the Cashier's Office.

¹⁰⁹See Appendix C for microphone frequency response.

A Bruel and Kjaer Presion Sound Level Meter was utilized to measure the exact sound levels at the teller's window and in the alcove.¹¹⁰ Readings were taken at two different times. The ranges of the sound levels for the teller's window were 64 to 80 db's (C scale); for the stand in the alcove, the range was 51 to 68 db's (C scale). 34 db's is the sound level in a library; 54 db's is the sound levels in a typical business office; 65 db's is the sound level of average conversational speech; 74 db's is the sound level of average street traffic; 88 db's is the sound level of the inside of a bus; and 94 db's is the average sound level inside a New York subway train.¹¹¹ As can be seen from these figures, the sound level of the alcove was substantially lower than the rather noisy level of the teller's window.

Standard Scotch Brand, 5 inch reel, magnetic tape (190 series) was used to record the speech of the volunteers.

It had been decided that an equal number of males and females plus some extra speakers, in case a recording was not good, should be obtained. In addition, permission had been obtained from the director of the Audiology and Speech Sciences Department's research project to state that the Audiology and Speech Sciences Department was the Department conducting the research. A sign was posted on the stand in the alcove and on the side of the teller's window being used stating that this was the voiceprint research project.

¹¹⁰Instructions and Application of the Precision Sound Level Meter
(Naerum, Denmark: Bruel and Kjaer, 1965).

¹¹¹Ibid., p. 4.

The sentences used were two:

Please give me my money; I want it.

My money is on the counter; please give it to me.

The procedures for the recordings were these. One technician was stationed in the teller's window. Another technician was stationed near the door of the Cashier's Office. As a possible speaker either came in or went out, the technician approached him and asked if he would like to participate in a research project. If the person accepted, a brief statement of the research goals and procedures was made. He was then led to the teller's window where the technician there recorded his name, age, nationality and state, if he was an American citizen, and local address. The technician also recorded the sex of the speaker and gave him a speaker number. The speaker's instructions were then given to him. He was to read, twice, the two sentences. No attempt was made to position him in front of the microphones. The speaker stated his speaker number and then the two sentences. When this was done, he was led to the alcove where he again stated his speaker number and repeated the two sentences twice. At this station, however, the speaker held the microphone himself approximately eight to twelve inches from his lips. The range of the distance of the speaker from the microphones in the teller's window was approximately 16 to 28 inches. Some speakers leaned on the counter, others stood back from it.

Eighty-four speakers in all were recorded, 42 males and 42 females. Most of the speakers were natives of lower Michigan. Almost all were undergraduates of 18 to 20 years of age.

Analysis of the Recordings

The recording sessions took two days. At the beginning and the end of each day, the tapes were marked for the date, place of recording, tape recorder used, and speakers on the tape. New tapes were used each day.

When the recordings were finished, the tapes were taken over to the Audiology and Speech Sciences Department. There, using a Voice-print Laboratories Sound Spectrograph, an assistant spent five weeks running off over 1,500 spectrograms. Parts of 16 speaker recordings were too poor to use. Those parts of these 16 which were useable were spectrographically analyzed and used in the identification tasks later. The resulting spectrograms of all the recordings were spectrograms of words in a fixed context. Bar spectrograms were made. The settings of the spectrograph were normal display, wideband filter, high shaping, linear expand frequency scale, and delay gate out. These are the normal settings for the making of bar spectrograms.

The resulting spectrograms then had the words marked underneath the sound representation and a five digit coded affixed to them. The five digit code number as before was from the Rand Corporation's book of 1,000,000 random digits.¹¹² The three field recorders had three distinct code sequence numbers to prevent an identifier from memorizing speaker numbers from seeing these numbers too often.

The identification tasks were then undertaken. A total of 300 such tasks were done. One identifier did 120; the other did 180. Each identifier was paid \$2.00 an hour.

¹¹²Rand Corporation, op. cit.

The procedures for the identification tasks were as follows. Each identifier worked separately. He was given 10 unknown spectrograms at a time. One known was given to him at a time. He was to determine whether or not the known was in the unknown pile. The identifier was told that these would be either all male or all female voices; that there would be matches and no-matches; that all the spectrograms were contemporary speech samples; and that there could be double matches in the unknown spectrograms. Each identifier had the use of a desk or a table, whichever he preferred. He was given as much time as he needed to reach a conclusion. When he finished one task, he would bring the known sample back and state his conclusion. He would then be told whether he was right or wrong. If wrong, the correct answer would be given him to allow him to compare the two matching samples. He would then be handed a new known.

Ten known tasks comprised one section of a trial. There were six sections to each trial: there were three recorders and two levels of the sex factor for each recorder. After each section, the identifier was given a ten minute rest period.

The reason for giving one known to compare with ten unknowns is that until a large file or library of voices is gathered, the procedure for comparison of spectrograms will be just this method. Unknown speakers will be available; as a person becomes a suspect, his voice will be checked out against the unknown speaker samples for possible identification.

Results

The results of the main study's identification tasks are listed in Table 6.

Table 6. Results of the Main Study
of the School of Criminal Justice.

<u>Identifier</u>	<u>Matches</u>	<u>No-Matches</u>	<u>Total</u>
First Identifier	60% (94/157)	69% (15/23)	61% (109/180)
Second Identifier	52% (54/103)	76% (13/17)	55% (67/120)
Total	57% (148/260)	70% (28/40)	59% (176/300)

Recorder Results, Totals

	<u>Total</u>	<u>Matches</u>
Uher	51%	(51/100)
Wollensak (omni. mike)	60%	(60/100)
Wollensak (uni. mike)	65%	(65/100)

Sex Results, Totals

	<u>Total</u>	<u>Matches</u>
Males	62%	(94/150)
Females	54%	(82/150)

A factor analysis (treatments by treatments by subjects) was carried out on the data. Neither the sex factor, the recorder factor, nor the microphone factor showed significant results. A Chi square analysis also showed insignificant results. However, a Chi square analysis did show that the accuracy levels of the identifiers were significant, as Table 7 indicates.

Table 7. Results of Chi Square Analysis of Data
from the Main Study of the School of Criminal
Justice. Levels of significance shown.

<u>Identifier</u>	<u>Matches</u>	<u>No-Matches</u>	<u>Total</u>
First Identifier	.01	.01	.01
Second Identifier	.01	.01	.01
Total	.01	.01	.01

The expected frequency was 1 in 8; that is, the identifier had a 1 in 8 chance to make a correct identification.¹¹³ The results show that each identifier's total of correct identifications was significant to the .01 level. Yet the overall accuracy rate was 59%.

There are really two results of this study, a practical one, and a scientific one. The practical result was the 59% accuracy rate. The scientific result was the .01 level of significance of the 59% accuracy rate. What does this mean? Basically, it means that the fact that the identifiers got as many right as they did was statistically significant. It also means that in a practical application, the error rate possibility is 43% that a person will be misidentified and 30% that an innocent man will be identified when there is not a match present.

Four reasons for these results are possible. First, the most obvious possible reason is that the technique of acoustic spectrography does not work. However, the high level of significance definitely refutes this. The other three reasons offer a more plausible explanation.

¹¹³Dr. Stapleton, Chairman of the Statistics and Probability Department, Michigan State University, personal interview, September 11, 1970.

The amount of noise shown in the unknown, field spectrograms was substantial. The noise is the second reason. For, as one of the aims of this study was to make the field recordings as real as possible, the recordings may have been too practical. While the situation was real, the amount of noise appearing on the spectrograms represented a difficult problem for identifications. For how is it possible to distinguish what was noise and what was the voice resonance, without listening to the tape recordings.

This question leads to the third possible reason, the methodology of the identification tasks. The identifier received only the spectrograms. He never heard the tapes. If an identifier could have heard the tapes, he could have heard the noise and distinguished it from the vocal formants and resonance spectra. However, because of the way an experiment of this type must be designed, an identifier is not allowed to listen to the tape and look at the spectrogram. The manner in which the recordings must be made, with a speaker number given to and spoken by each speaker, means that the identifiers may not listen to the recordings. He has only the spectrograms to examine.

The fourth reason has to do with the training of an identifier. Kersta offers a two week training course; the identifiers in this project underwent as much as six weeks of training. What is needed for a truly comprehensive training program to qualify an individual to be an identifier is more than either Kersta's training course or the six week course taken by one identifier in this project. No attempt is made here to propose a full-scale program, but, nevertheless, some of the more obvious problems to be solved can be enumerated. First, more training on the spectrographic appearance of phonemes must be given.

Second, much more experience in identifying, and using tapes along with the spectrograms, is needed. Third, more than just two fixed context sentences must be used. A good identifier must have at least a working knowledge of what to expect from phonemes in random contexts.

Audiology and Speech Sciences' Results

The initial results of the Audiology Department's project have been considerably higher.¹¹⁴ Their mean total for all identifications of the type of spectrograms used in the School of Criminal Justice's project was 88%.¹¹⁵ The methodology of the Audiology Department does differ significantly from the research design of the School of Criminal Justice.¹¹⁶ Audiology does use matching-to-sample type tasks as well as true identification tasks in their procedures. In addition, they did not encounter as high amplitude background noise as did this research project.

Nevertheless, because there was such a large discrepancy between the two projects' results, a number of spectrograms from the School of Criminal Justice's project were brought over for tests with the identifiers of the Audiology Department's project. The procedure followed was the matching-to-sample method. The initial results indicate that the second, third, and fourth reasons offered in explanation for the low percentage of accuracy may indeed have some truth in them.

¹¹⁴"Voice Identification Project," op. cit.

¹¹⁵Dr. Oscar Tosi, personal interview, op. cit.

¹¹⁶Voice Identification Proposal, op. cit., Exhibit A, pp. 3-28.

Out of 144 tasks, 50 correct matches were made. This is a 35% accuracy level.¹¹⁷ It would definitely seem that the background noise, lack of ability to listen to the tape recordings of the speech and the noise, and lack of experience with this type of problem resulted in an inability to abstract the pertinent information from the spectrogram to make a correct identification. It is really significant that the identifiers of the School of Criminal Justice did so well.

CHAPTER SUMMARY

Two studies were carried out by the researchers of the School of Criminal Justice. The results of the pilot study showed a 70% accuracy rate. The results of the main study showed a 59% accuracy rate. The factors of different recorders and male and female vocal cord differences proved to have insignificant effects on the accuracy rate.

It is hypothesized that there are three main causes for the relatively low rate of the main study. First, the amount of background noise obscured the essential spectrographic cues for correct identifications. Secondly, the methodology of the research prohibited the identifiers from listening to the recordings of the speakers and thus learning which resonance patterns were noise and which were formant patterns. Thirdly, the training of the identifiers did not prepare them adequately for the problems they encountered in the identification tasks.

The technique is considered to be feasible. The fact that more experimentation is needed before it is definitely concluded to be valid and reliable is obvious.

¹¹⁷Results furnished by Mrs. Van Huss, Assistant Researcher, Audiology and Speech Sciences Department Voiceprint Research Project.

Chapter V

THE LEGAL DILEMMA OF VOICEPRINT

Voiceprint identification has not received complete scientific acceptance. It has not fared much better in the courts. Scientific and legal judgments differ in the respect that scientific acceptance is based on technical evidence; court acceptance relies heavily on the opinions of expert witnesses.¹¹⁸ The distinction between the scientific and legal acceptance is real; yet, as can be seen, legal acceptance does rest upon scientific acceptance, and ultimately on scientific proof. Nevertheless, when experts differ in their opinions, the court may leave the assessment and final determination of the worth of conflicting opinions and the relative expertise of the witnesses to the jury.¹¹⁹ Voiceprint identification has been used in five major cases. In four of these cases, it has been used as evidence for the prosecution. In one case, it was used as evidence for the defense. These five cases are presented in this chapter. The question of the admissibility of voiceprint identification is also dealt with, as well as the Constitutional aspects of voiceprint identification.

¹¹⁸Bolt, op. cit., 13.

¹¹⁹Ibid., 13.

VOICEPRINT IN COURT

United States v. Wright

The only case to date in which voiceprint identification has been admitted as evidence for the prosecution and which has been affirmed by a higher court is the United States v. Wright.¹²⁰ This case was decided by a military court, and the conviction of Wright was confirmed by a military court of appeals. In this case, the Court reasoned that the testimony of experts has been consistently recognized in areas in which there is neither infallibility of result nor unanimity of opinion as to the existence or non-existence of a particular fact or condition. Three examples were cited. First, the difference of opinion among psychiatrists as to the mental condition of a particular defendant was cited. Second, handwriting analysis of a questioned document was specified. Third, the admissibility of aural identification of a person's voice was cited.

"Since voice identification by ear is fully acceptable in the courts, the court members could best determine for themselves the margin of error, if any, in Mr. Kersta's expert opinion."¹²¹

People v. Straehle

In 1966,¹²² the New Rochelle, New York, police department was attempting to obtain evidence against a local bookmaker. However,

¹²⁰"U.S. v. Wright," Criminal Law Reporter (August 23, 1967), 2287.

¹²¹Ibid., 2287.

¹²²Oscar Tosi, "Speaker Identification Through Acoustic Spectrography," op. cit., 26.

whenever they conducted a raid on a location suspected to be his center of operations, no trace of any illegal gambling activities could be found. Because of the circumstances of the case, phone taps were placed on the telephones of the police department to determine if a police officer was warning the bookmaker. A call by a policeman from the headquarters building was recorded in which a man warned the person, supposedly the bookmaker, of an impending raid on his center of bookmaking operations. Officer George Straehle was the suspected policeman. Mr. Lawrence G. Kersta was contacted and requested to use his technique of acoustic spectrography to determine the identity of the policeman who had warned the bookmaker. Kersta identified Straehle.

Evidence of Kersta's identification was admitted in the subsequent court case. The ruling on the admissibility of Kersta's expert opinion was that the admissibility of the voiceprint evidence was solely dependent upon whether or not the sound spectrograph qualified as a scientifically accurate instrument. The testimony of Kersta was thus admitted. The jury had the responsibility of weighing the evidence in their deliberations. The result of the case was a hung jury.

California v. King

Edward Lee King was convicted in 1966¹²³ on one count of arson which occurred during the Watts riot. During the trial, Kersta testified that he had matched a sample of King's voice to the voice of a man who had appeared on television and claimed he had set fire to five

¹²³"California v. King," Criminal Law Reporter (October 30, 1968), 2086-87.

businesses during the Watts riot. On the basis of this and other evidence, the jury convicted King on one of the counts of arson with which he was charged.

However, the Second California Court of Appeals in 1968 reversed and remanded the case for retrial. In reversing the conviction, the Court stated that voiceprint has not been scientifically accepted and cannot be accepted by the Court.

"Kersta's admissions that his process is entirely subjective and founded on his opinion alone without general acceptance within the scientific community compels us to rule that 'voiceprint' identification process has not reached a sufficient level of scientific certainty to be accepted as identification evidence in cases where the life or liberty of a defendant may be at stake."¹²⁴

New Jersey v. Di Gilio

This case¹²⁵ involved extortion demands upon Mr. Julius Pereira. Pereira was the owner of Pu-Rite Car Wash, in Woodbridge, New Jersey. In 1966 and 1967, he had borrowed \$2,000 from a man named Di Gilio, a loanshark. From 1967 to 1970, Pereira had been paying Di Gilio back. Finally, in 1970, Pereira had already paid Di Gilio \$6,500 on the \$2,000 loan. Pereira then received threatening phone calls of which one was recorded. On the phone, the caller yelled, "I'll come down and chop your ---- head off."¹²⁶ The police arrested Di Gilio.

¹²⁴Ibid., 2086

¹²⁵Reginal Kavanagh, "Voiceprint Rules Out Di Gilio as Maker of Phone Threat," The Daily Home News (New Brunswick, New Jersey, March 7, 1970), 1.

¹²⁶Ibid., 1.

They took a sample of his voice and sent it to Kersta with the recording of the telephone call. Kersta, Dr. Oscar Tosi of Michigan State University, and Dr. Louis Gerstman, a psychology professor and speech researcher from New York University, were allowed to testify because the witnesses' purposes in using the technique of voiceprint identification were to exclude the defendant rather than identify him. Di Gilio was found innocent of the charge of extortion.

New Jersey v. Cary

This case¹²⁷ is the most important case to date for voiceprint identification. The importance stems not so much from the details of the case, but from the legal course the case followed.

The case first came before a Superior Court in the State of New Jersey. The Court was requested by the prosecution to allow an exemplar of the defendant's voice to be taken for comparison by voiceprint identification. The Superior Court granted this request on the grounds that a person's voice is non-testimonial physical evidence and is not protected by the Fifth Amendment's self-incrimination clause. The defense, though, appealed this ruling to the New Jersey Supreme Court.

The Supreme Court of New Jersey handed down this ruling:

"We believe that before an intrusion into a person's privacy can be proper within the protections afforded by the Fourth Amendment, the product of the search must have the capacity to be admissible in evidence."¹²⁸

¹²⁷"New Jersey v. Cary," Criminal Law Reporter (March 20, 1968), 2485-86.

¹²⁸"New Jersey v. Cary," Criminal Law Reporter (June 9, 1967), 2181.

The Supreme Court ruled that a person's voice is non-testimonial physical evidence and is proper evidence. Aural recognition evidence is proper and admissible evidence. However, relying on Warden v. Hayden¹²⁹ the New Jersey Supreme Court ruled that the Fourth Amendment was controlling. The right to privacy demanded that if the right were to be forfeited, good reason must be supplied. Due to the untested nature of voiceprint, the Supreme Court refused to compel the defendant to submit to a recording of his voice. They further ruled that, while a trial judge does not normally rule on the admissibility of evidence until it is presented at trial, the trial judge must make some determination whether or not the results of such a test as voiceprint would produce evidence which would be admissible at trial. The right to privacy cannot be invaded by a search unless admissible evidence is known to be a potential result. It is because the validity and reliability of voiceprint identification was in question that the Supreme Court ordered the trial judge to make a determination as to the admissibility of voiceprint identification in a pre-trial hearing before ordering a voice exemplar to be given. The case was remanded for further testimony.

In the pre-trial hearing,¹³⁰ the trial judge noted that where there is no general scientific acceptance of the reliability of a new scientific process, it cannot be taken under Judicial Notice.

¹²⁹387 U.S. 294, 87 Sup. Ct. 1642, 18 L. Ed. 2d 782 (May 29, 1970).

¹³⁰"New Jersey v. Cary," (March 20, 1968), op. cit., 2485-86.

"When scientific aids to the discovery of truth receive general recognition scientifically as to their accuracy, courts will take judicial notice of this fact and admit evidence obtained through their use. We do not believe that the spectrograph presently meets that test. We are satisfied that the spectrograph is an efficient and accurate piece of equipment which produces an accurate spectrogram from the tape recording. But whether it produces all the detail required to provide reasonably certain individualistic identification appears to be scientifically questioned at the present period."¹³¹

This, then, is exactly where voiceprint identification stands: until it gains scientific acceptance, legal acceptance will be withheld. The legal admissibility of voiceprint identification is dependent on the scientific determination of its validity and reliability.

GENERAL ADMISSIBILITY REQUIREMENTS

A general admissibility requirement for identity tests is that the significant features which appear on the compared objects must not recur so frequently on other objects that uncertainty results as to whether or not the two objects being compared are actually identical. A mathematical expression of this requirement is the probability of the re-appearance of the particular combination of identifying characteristics on non-identical samples.¹³²

The Frye Rule

"While the courts will go a long way in admitting expert testimony, deduced from a well-recognized scientific principle or discovery, the thing from which the

¹³¹"New Jersey v. Cary," Criminal Law Reporter (February 26, 1968), 2486.

¹³²Kamine, op. cit., 36.

deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs."¹³³

This rule arose out of Frye v. United States, 1923, on the question of the validity and reliability of a technique which claimed to be able to tell whether a person was telling the truth or not. The Frye Rule is the rule used to determine if a new scientific aid meets the requirements of admissibility; it must have gained general acceptance in the field of science in which it arose.

This rule was used in the King case and the Cary case.

In the 1949 Michigan case of the People v. Morse,¹³⁴ a similar conclusion concerning the Harger Drunkometer was reached. The court ruled that until reasonable certainty can follow from scientific tests, demonstrable by experts in the giving of these tests, the court is not in error when it refuses to admit results of tests that do not demonstrate reasonable certainty. It was granted that there is great difficulty in determining when a scientific principle or discovery crosses the line between unproven theory and fact. But, until such a demonstration is made, the evidential force of the testimony of experts about the new principle or discovery will not be accepted by the courts. Until the principle gains general acceptance in the particular field in which it belongs, it is not admissible as evidence.

¹³³Frye v. U.S., 293 F. 1013, 34 A.L.R. 145 (D.C. Cir., 1923), p. 1013.

¹³⁴People v. Morse, 325 Mich. 270, 274-5; 38 N.W. 2d 322 (1949), p. 1092.

Firearms identification,¹³⁵ fingerprints and photographs of fingerprints,¹³⁶ footprints,¹³⁷ ballistics tests,¹³⁸ Nalline tests,¹³⁹ and Drunkometer tests and blood tests¹⁴⁰ all met the requirements of the Frye Rule. Generally, all are admissible.

Yet, the Frye Rule is not the only rule that is followed. In New Hampshire v. Roberts in 1960,¹⁴¹ the trial court ruled that evidence does not have to be infallible to be admissible. It was ruled that evidence was an aid to judge and jury; its deficiencies or weaknesses were matters for the defense to point out. These deficiencies affect the weight of the evidence but do not determine the admissibility of it. This was the rule used in the case of United States v. Wright.

Voiceprint identification does not yet meet the Frye Rule. Until a comprehensive and scientifically acceptable experiment is carried out and shows scientifically and practically acceptable results, voiceprint will continue to be rejected as invalid and unreliable by both science and law. The experiment of the School of Criminal Justice has laid some foundations for further research. However, the research project

¹³⁵Evans v. Commonwealth, 230 Ky 411, 19 S.W. 2d 1091 (1929), p. 1092.

¹³⁶People v. Jennings, 252 Illinois 534, 96 N.W. 1077 (1911), p. 1077.

¹³⁷Elmore v. Commonwealth, 282 Ky. 443, 138 S.W. 2d 956 (1966), p. 956.

¹³⁸People v. Fisher, 340 Ill. 216, 172 N.E. 743 (1930), p. 744.

¹³⁹People v. Williams, 164 Cal App. 2d 858, 331 R. 2d 251 (1958), p. 251.

¹⁴⁰Schmerber v. California, 384 U.S. 757 (1965), p. 757.

¹⁴¹State v. Roberts, 102 N. H. 414, 158 A. 2d. 458 (1960), p. 459.

of the Audiology and Speech Sciences Department does promise to fulfill at least some of the requirements for voiceprint acceptance.

If voiceprint identification or a modified form of the technique of acoustic spectrography does become scientifically recognized and admissible as evidence in court, in what relation does it stand to the constitutional requirements of the Fourth, Fifth, and Sixth Amendments?

VOICEPRINT IDENTIFICATION AND CONSTITUTIONAL LAW

This section is not a totally comprehensive review of all cases relating to the constitutional areas covered by the Fourth, Fifth, and Sixth Amendments of the Constitution of the United States. The most recent cases relating to the three areas most critically involved with voiceprint identification are covered. These three areas deal with the right to privacy and protection from illegal searches and seizures, the Fourth Amendment; the privilege against self-incrimination, the Fifth Amendment; and the right to assistance of counsel, the Sixth Amendment.

Fourth Amendment Law

In Katz v. United States, in 1967,¹⁴² the protection of the Fourth Amendment was judged to protect people, not places. The right to privacy and protection from illegal searches and seizures extends not only to a person's home, for example, but also to the person himself. A person's body or mind cannot be invaded, searched, or seized. This principle had been emphasized before in the Rochin case.¹⁴³ In the

¹⁴²Katz v. U.S., 389 U.S. 347 (1967), p. 350.

¹⁴³Rochin v. California, 342 U.S. 165, (1952), p. 165.

Rochin case, a suspect's stomach was pumped to recover some pills he had swallowed which were suspected by the police to be illegal narcotics. This invasion of a person's body led to the "shock the conscience" rule: an invasion of a person's body by such unreasonable techniques, based on mere suspicion, are violations of a person's right to privacy and protection from illegal search and seizure.

The application to voiceprint has already been referred to in the Cary case. What is meant with regards to voiceprint is that there must be strong grounds for requesting a voice sample and comparison of this exemplar; the results of any such test must be potentially admissible as evidence before the individual's right to privacy is invaded by the police.

Fifth Amendment Law

The rationale behind this judgment is that, just as a person's home is part of his privacy, so too is his body. For evidence obtained from a person's body falls in the category of non-testimonial, physical evidence; that is, the voice of an individual is physical evidence, just as his fingerprints are. And just as fingerprints, his voice does not in itself represent testimonial evidence. Warden v. Hayden, 1967, established that while the right to privacy must be protected, non-testimonial, physical evidence does not fall under the protection of the Fifth Amendment; non-testimonial, physical evidence is not protected by the self-incrimination clause of the Fifth Amendment. Schmerber v. California in 1965,¹⁴⁴ had emphasized the distinction between testimonial and non-testimonial evidence. In Schmerber,

¹⁴⁴Schmerber v. California, op. cit., 757.

an involuntary blood test was allowed to be admitted into evidence; it was not rejected as a violation of the defendant's right to protection from forced self-incrimination. Further, in United States v. Wade, 1967,¹⁴⁵ Justice Brennan stated that the defendant in a line-up can be compelled to repeat the exact words spoken at the scene of a crime. The voice is physical evidence of a non-testimonial nature and is not protected by the Fifth Amendment.

Sixth Amendment Law

Finally, the assistance of counsel is recommended whenever a voice sample is taken for voiceprint identification.¹⁴⁶ While the presence of counsel is not required, the judgment in Gilbert v. California¹⁴⁷ strongly recommended it. For the Fifth Amendment Rights of the defendant must be explained so that the defendant realized that he must give a voice exemplar, that his voice is not protected by the Fifth Amendment's right to silence.¹⁴⁸ In addition, the sample taken should not necessarily be the same words spoken as the unknown voice. For, if a comparison were made of the defendant's voice speaking the same incriminating message as the unknown voice, prejudice in the comparison might be claimed. Finally, presence of counsel may be required for the taking and comparing of a voice exemplar if these actions are classified as a line-up type procedure.¹⁴⁹

¹⁴⁵Wade v. U.S., 388 U.S. 218 (1967), p. 218.

¹⁴⁶Ibid., p. 218.

¹⁴⁷Gilbert v. California, 388 U.S. 263, (1967), p. 263.

¹⁴⁸Miranda v. Arizona, 384 U.S. 463 (1966), p. 436.

¹⁴⁹Wade v. U.S., op. cit., p. 218.

In general, the constitutional requirements and protections do not appear to present any difficulties for voiceprint identification; in fact, they aid the voiceprint process.¹⁵⁰

CHAPTER SUMMARY

Neither the scientific community nor the courts of law have accepted voiceprint identification. The basis for recognition of this technique is scientific proof of it's reliability and validity. Until this proof is presented, voiceprint identification remains in scientific and legal disfavor. When voiceprint identification is finally proved to be scientifically acceptable and is admitted as evidence in court, the requirements of the Fourth, Fifth, and Sixth Amendments of the Constitution of the United States must be met for voiceprint to be acceptable on constitutional grounds.

¹⁵⁰R. Steinhauer, "Voice Prints," Saturday Review, 52, September 6, 1969), 58.

Chapter VI

SUMMARY AND CONCLUSIONS

In the preceding chapters of this study, the history, the theoretical background of voiceprint identification, the experimental research carried out concerning voiceprint, the recent research of the School of Criminal Justice on voiceprint identification, and the admissibility of voiceprint identification as evidence in court have been presented and discussed.

The theory behind voiceprint identification extends from the very basic concepts of sound, resonance, and sound waves through the human production of sound in speech and the phonetic classification of sounds to the visible representations of speech as portrayed by the sound spectrograph on spectrograms. The most important area of the theoretical foundations of voiceprint is the theory of invariant speech. If each individual's voice is unique and the features that make each person's voice unique can be portrayed accurately by the sound spectrograph, then voiceprint identification is, indeed, a feasible method of identification of suspects by their voices.

Nevertheless, the "if's" must be proven; and the experimental literature to date has neither proven nor disproven that voiceprint identification is a valid and reliable identification technique. Lawrence Kersta's experiments have shown highly promising results;¹⁵¹

¹⁵¹Kersta, "Voiceprint Identification," Nature, op. cit., 1253-57.

at the same time, his experimental procedures have been heavily criticized by other researchers. Researchers such as Stevens, Williams, Carbonell, and Woods¹⁵² have produced data which shows that aural recognition is superior to the voiceprint technique. The original research of the School of Criminal Justice, while permitting partial conclusions to be drawn, was not intended as nor did it result in a definitive answer to the questions of the validity and reliability of voiceprint.

The research project conducted by the School of Criminal Justice was designed to lay the foundations for further research. Practical field tape recordings were made in a functioning bank. Spectrograms of the tape recordings were made on a Voiceprint Laboratories Sound Spectrograph. These spectrograms were used in identification tasks with the results of a 59% accuracy rate, significant to the .01 level. This research identified certain aspects of the voiceprint identification process which indicate future valuable research areas as further discussed in the proposals section of this chapter. This experiment indicates that this process is feasible. x

The problem of the admissibility of voiceprint identification, though, does not rest on the concept of feasibility. Legal acceptance of voiceprint as an identification method rests on experimental evidence of the worth of voiceprint as required in the Frye Rule.

¹⁵²Stevens, op. cit., 1596-1607.

CONCLUSIONS

What, then, can be concluded from this study? Five basic conclusions may be drawn from the information contained in this study. The conclusions must be based on the results of the experiments presented in Chapter III, the results of the experiment conducted by the School of Criminal Justice, and the results of the court cases, reported in this thesis, in which voiceprint identification has appeared.

(1) The experiments reported in this paper clearly illustrate the need for adequate and careful training of those engaged in voice identification through the technique of voiceprint identification.

(2) Experiments conducted under field conditions with equipment of limited potential (technical excellence) produced raw data (spectrograms) that was difficult to analyze.

(3) The relatively small number of field tests (300 separate identifications using approximately 1500 spectrograms from 84 different speakers) yielded the rather low accuracy rate (correct identification) of 59%; however, a chi-square analysis of the data indicates an acceptable significance level of .01.

(4) It is the opinion of this writer that the technique of voiceprint identification is a feasible identification method. Results obtained from carefully controlled experimental situations are impressive with regard to accuracy. The relatively low accuracy percentage rate is not discouraging.

(5) It is apparent that if the technique is to be used in conjunction with conventional law enforcement investigative techniques:

(a) proper training of identifiers is of the utmost importance.

(b) additional experimentation must be conducted to develop a method whereby field recordings will produce better quality spectrograms.

Better training, better equipment, a larger and more comprehensive study, all these factors must be incorporated in future research. The technique is feasible.¹⁵³ Nevertheless, further research is necessary; and this research must be of the highest caliber. In addition, in a real application of voiceprint identification, the identifiers must have undergone a substantial amount and a stringent period of training. Poorly trained identifiers can only hurt the case of voiceprint in court.

PROPOSALS

In view of the previous conclusions, proposals for further research in the area of voiceprint identification are necessary. Specifically, research needs to be carried out in four main areas:

(1) The training of identifiers must be studied in order to determine what they must learn, how they should be taught, and when they are qualified to make identifications.

(2) In any experiment to be carried out in the future, the size of the population sample must be increased to include a much larger number of speakers. The size could range from 200 to 1,000 and beyond.

¹⁵³Kersta has used a maximum of 123 voices. Kersta, "Voiceprint Identification," Newsletter, op. cit., 25.

In addition, the procedures for any identification tasks must take the size of the possible real population and possible practical field applications into consideration.

(3) The recording of the sample voices must be done in field situations. Laboratory recordings will not be acceptable as proof. For the environment in which the recordings are made does have an effect on what is recorded. The second aspect of this point is that the manner of recording must be rethought. A technique must be identified that would allow the pertinent speech sounds to be recorded and the irrelevant noise elements to be blocked out.

(4) The actual speech sounds that are recorded should be made of the same phonemes, but of different words, phrases, and sentences. The School of Criminal Justice's sentences by design resulted in the entire spectrogram being a relatively steady-state whole. All the words and sounds of all the speakers were of the same phonemes spoken in the same order. A real application of voiceprint identification would not have this advantage.

BIBLIOGRAPHY

A. BOOKS

- Abercrombe, David. Elements of General Phonetics. Chicago: Aldine Publishing Company, 1967.
- Beranek, Leo. Acoustics. New York: McGraw-Hill, 1964.
- Black, John. Speech. New York: McGraw-Hill, 1955.
- Bruel & Kjaer. Instructions and Applications: Precision Sound Level Meter. Denmark: Bruel & Kjaer, 1965.
- Carrell, James. Phonetics. New York: McGraw-Hill, 1960.
- Cummins, H., and Midlo, C. Fingerprints, Palms, and Soles: An Introduction to Dermatoglyphics. New York: Dover Publications, 1961.
- Dictionary of Electronic Terms (Chicago: Allied Radio, 1959).
- DuBrul, E. Speech Apparatus. Springfield, Illinois: Thomas Publisher, 1958.
- Flanagan, James. Speech Analysis, Synthesis, and Perception. New York: Springer Verlag, 1965.
- Fletcher, Harvey. Speech and Hearing in Communication. New York: Von Nostrand, 1953.
- Goldstein, Max. The Acoustical Method for the Training of the Deaf. St. Louis: Laryngoscope Press, 1939.
- Hancock, John. Communication Theory. New York: McGraw-Hill, 1961.
- Hans, Erni, ed., Communication and Language. London: MacDonald Company, 1965.
- Harrah, David. Communication. Cambridge, Massachusetts: Massachusetts Institute of Technology Press, 1963.
- Hoops, Richard. Acoustics in Speech. Springfield, Illinois: Thomas Publisher, 1960.
- Joos, Martin. Acoustic Phonetics. Baltimore: Linguistic Society of America, 1948.
- Judson, Lyman S., and Weaver, Andrew T. Speech Science. New York: Appleton-Century-Crofts, 1965.
- Kaplan, H. Anatomy and Physiology of Speech. New York: McGraw-Hill, 1960.

- Ladefoged, Peter. Elements of Acoustic Phonetics. Chicago: University of Chicago Press, 1962.
- _____. Three Areas of Experimental Phonetics. London: Oxford University Publishers, 1967.
- Lehiste, Leo. Acoustic Phonetics. Cambridge, Massachusetts: Massachusetts Institute of Technology Press, 1967.
- Mackenzie, George. Acoustics. New York: Focal Press, 1964.
- Malmberg, Bertil. Manual of Phonetics. Amsterdam: North Holland Publishing Company, 1968.
- McCormick, Charles. Law of Evidence. St. Paul: West Publications Company, 1954.
- McMullen, Charles. Communication Theory Principles. New York:
- O'Neill, Edward. Communication and Information Theory Aspects of Modern Optics. Syracuse, New York: General Electric Company Company Electronics Laboratory, 1962.
- Potter, Ralph, Kopp, G., and Kopp, H. Visible Speech. New York: Dover Publications, Inc., 1947.
- Pulgram, E. Introduction to the Spectrography of Speech. 's Gravenhage: Mouton, 1959.
- Radio Builder's Handbook. Chicago: Allied Radio, 1959.
- Rand Corporation. A Million Random Digits with 100,000 Normal Deviates. Glencoe, Illinois: Free Press Publishers, 1955.
- Richardson, Edward. Sound. London: Arnold Publisher, 1953.
- Shearer, William M. Illustrated Speech Anatomy. Springfield, Illinois: Thomas Publisher, 1968.
- "Sound," Collegiate Encyclopedia (1970), XVII, 301-304.
- Stewart, George, and Lindsay, Robert. Acoustics. New York: Van Nostrand, 1930.
- Turner, William. Invisible Witness. New York: Bobbs-Merrill Company, Inc., 1968.
- Voiceprint Laboratories. Operating Manual: Sound Spectrograph Model Model 4691 A. Somerville, New Jersey: Voiceprint Laboratories, 1967.
- West, Robert. Phonetics. New York: Harper Publications, 1941.

Weston, Paul, and Wells, Kenneth. Criminal Investigation: Basic Perspectives. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

Wigmore. Wigmore's Code of the Rules of Evidence. Boston: Little, Brown and Company, 1942.

Wise, Claude. Phonetics. Englewood Cliffs, New Jersey: Prentice-Hall, 1957.

Wood, Alexander. Acoustics. New York: Dover Publications, 1960.

Zemlin, Willard. Speech and Hearing Science. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

B. PERIODICALS

"Airman Convicted on Voiceprint Test," Crime Control Digest, August, 1967, 16.

Berry, J. "Voiceprints: Poison for the Telephone Rat," Popular Science, 167 (September, 1965), 80-83.

Bolt, Richard, and others. "Identification of Speakers by Speech Spectrograms," Science, 166 (October 17, 1969), 338-343.

"Break in the Greatest Story of Newspaper History," Newsweek, 5 (February 23, 1935), 23-24.

Bricker, Peter, and Pruzansky, Sandra. "Effects of Stimulus Content and Duration on Talker Identification," Journal of the Acoustical Society of America, 40 (August, 1966), 1441-49.

Carvin, Paul and Ladefoged, Peter. "Speaker Identification and Message Recognition in Speech Recognition," Phonetica, 9 (1963).

Clarke, F. "Speaker Recognition by Humans," Journal of the Acoustical Society of America, 37 (August, 1965), 1211.

Clarke, F., and Becker, R. "Comparison Techniques for Discrimination Among Talkers," Journal of Speech and Hearing Research, 12 (1969), 747-761.

Coon, Thomas. "Voiceprint Identification," Police (May-June, 1966), 90-92.

_____. "Voiceprint Identification Goes to Work," Police, 11 (September-October, 1966), 67.

"California's First Conviction Obtained on the Basis of Voiceprints," Crime Control Digest, October, 1968, 4.

- Dickson, D., "An Acoustic Study of Nasality," Journal of Speech and Hearing Research, 5 (1962), 103-111.
- "Doing New Tricks with Sound: Voiceprints," Business World, June 2, 1962, 80.
- Dudley, Homer, and Gruenz, Otto. "Visible Speech Translators with External Phosphors," Journal of the Acoustical Society of America, 18 (1946), 62.
- Fant, Gunnar. "Modern Instruments and Methods for Acoustic Studies of Speech," Proceedings of the Eighth International Congress of Linguistics, 1958, 282-358.
- _____. "Sound Spectrography," Proceedings of the Fourth International Congress of Phonetic Science, 1962b, 13-44.
- _____. "Descriptive Analysis of the Acoustic Aspects of Speech," Logos, 5 (April, 1962), 3-17.
- George. "Scientific Investigation and Defendant's Rights," 57 Michigan Law Review, 37 (1958), 39.
- Glenn, J., and Kleiner, N. "Speaker Identification Based on Nasal Phonation," Journal of the Acoustical Society of America, 43 (1968), 368-372.
- Liabersbrunner, H., Sebald, O., and Hantsche, H. "Zur Personfeststellung Mittels Stimmenund Sprachanalyse," Archiv Fur Kriminologie, July-August, 1963, 1-9.
- Hargreaves, W., and Starkweather, J. "Recognition of Speaker Identity," Language and Speech, 6 (1963), 63-67.
- "Heart Doctors Heed Telltale Voice; Coronary-Prone Patter in Voice," Business World, May 24, 1969, 130.
- Hill, T. "How Science Can Map Your Voice," Science Digest, 64 (November, 1968), 28-32.
- Johnson, J. "A Cathode-Ray Tube for Viewing Continuous Patterns," Journal of the Acoustical Society of America, 17 (November, 1946), 91.
- Kersta, L. "Amplitude Cross-Section Representation with the Sound Spectrograph," Journal of the Acoustical Society of America, 20 (November, 1948), 796.
- _____. "Voiceprint Identification," Nature, 196 (December 29, 1962), 1253-57.
- _____. "Speaker Recognition and Identification by Voiceprints," 40 Connecticut Bar Journal, 586 (1966), 592.

- _____. "Spectrographic Analysis of Body Sounds," Journal of the Association for the advancement of Medical Instrumentation, November-December 1966, pp. 7-10.
- _____. "Voiceprint Identification and Application," Fingerprint and Identification Magazine, May, 1970, 3-8 and 22.
- _____. "Voiceprint Identification," Newsletter of the American Academy of Forensic Sciences, June-July, 1970, 25-32.
- Keith-Smith, J. "Decision-Theoretic Speaker Recognizer," Journal of the Acoustical Society of America, 34 (August, 1965), 1908.
- Koenig, W., Dunn, K., and Lacy, L. "The Sound Spectrograph," Journal of the Acoustic Society of America, 17 (July, 1946), 19.
- Koenig, W. and Ruppel, A. "Quantitative Amplitude Representation in Sound Spectrograms," Journal of the Acoustical Society of America, 20 (November, 1948), 787-795.
- Kopp, G. and Green, H. "Basic Phonetic Principles of Visible Speech," Journal of the Acoustical Society of America, 18 (July, 1946), 74.
- Ladefoged, Peter and Broadbent, D. "Information Conveyed by Vowels," Journal of the Acoustical Society of America, 29 (1957), 98-104.
- Ladefoged, Peter, and Vanderslice, Ralph. "The 'Voiceprint' Mystique," Working Papers in Phonetics, 7 (November, 1967).
- Levine, A. "They Listen with Their Eyes," Science Digest, 21 (February, 1947), 85-87.
- Li, K., Damman, J., and Chapman, W. "Experimental Studies in Speaker Verification Using an Adaptive System," Journal of the Acoustical Society of America, 40 (1966), 966-978.
- Lieberman, A., and others. American Annals of the Deaf, 113 (1968), 127.
- Mann, M. "Your Voice Gives You Away: Voiceprints," Popular Science, 103 (August, 1962), 13.
- McDade, T. "The Voiceprint," The Criminologist, 7 (February, 1968), 52-60.
- Miller, R. "Nature of the Vocal Cord Wave," Journal of the Acoustical Society of America, 31 (1959), 667-677.
- More, Harry. "Voice Print - A New Scientific Aid for Law Enforcement," National Sheriff, March-April, 1969, 12-13, 26-28.
- "New Jersey V. Cary," Criminal Law Reporter, 1 (July 5, 1967), 2181.
- "People v. King," Criminal Law Reporter, 4 (October 30, 1968), 2085-2086.

- Peterson, Gordon. "Design of Visible Speech Devices," Journal of the Acoustical Society of America, 26 (May, 1954), 406-413.
- Peterson G. and Barney, H. "Control Methods Used in the Study of Vowels," Journal of the Acoustical Society of America, 24 (1952), 175-184.
- Pollack, I., Pickett, J., and Sumby, W. "On the Identification of Speakers by Voice," Journal of the Acoustical Society of America, 26 (1954), 403-406.
- Potter, R. "Visible Patterns of Sound," Science, 1 (November 9, 1945), 463-470.
- _____. "Introduction to Discussion of Sound Portrayal," Journal of the Acoustical Society of America, 18 (July, 1946), 4.
- Potter, R. and Steinberg, J. "Toward the Specification of Speech," Journal of the Acoustical Society of America, 22 (1950), 812-818.
- Presti, A. "High Speed Sound Spectrograph," Journal of the Acoustical Society of America, 40 (1966), 628.
- Prestigiacomo, A. "Plastic Tape Sound Spectrograph," Speech and Hearing Disorders, 22 (1957), 321-327.
- _____. "Amplitude Contour Display of Sound Spectrograms," Journal of the Acoustical Society of America, 34 (November, 1962), 1684.
- Pruzansky, S. "Pattern-Matching Procedure for Automatic Talker Recognition," Journal of the Acoustical Society of America, 35 (1963), 354-358.
- Pruzansky, S. and Mathews, M. "Talker-Recognition Procedure Based on Analysis of Variance," Journal of the Acoustical Society of America, 36 (1964), 2041-2047.
- Ramishivli, G. "Automatic Voice Recognition," Engineering Cybernetics, 5 (September-October, 1966), 84-90.
- Reisz, R. and Schott, L. "The Visible Speech Cathode-Ray Translator," Journal of the Acoustical Society of America, 18 (July, 1946), 50.
- Schroeder, M. "Vocoders: Analysis and Synthesis of Speech," Proceedings of the Institute of Electrical and Electronics Engineers, 54 (May, 1966), 720-734.
- "Seeing is Hearing: Sound Made Visible," Newsweek, 26 (November 19, 1945), 96.
- "Seeing Sound," Newsweek, 36 (July 3, 1950), 42.
- "Sound Judgement," Time, 89 (June 23, 1967), 66.

- "State v. Cary," Criminal Law Reporter, 2 (March 20, 1968), 2485-2486.
- Steinberg, J. and French, N. "The Portrayal of Visible Speech," Journal of the Acoustical Society of America, 18 (July, 1946), 4.
- Steinhauer, R. "Voice Prints," Saturday Review, 52 (September 6, 1969), 56-59.
- Stevens, K. and others. "Speaker Authentication and Identification: A Comparison of Spectrographic and Auditory Presentations of Speech Material," Journal of the Acoustical Society of America, 44 (1968), 1596-1607.
- Tall, J. "Voiceprinting and Its Uses," Saturday Review, 39 (September 29, 1956), 50-51.
- Tinker, F. "Thousand Ears of Soundprints," Popular Mechanics, 129 (June, 1966), 78-81.
- "Trial by Voiceprint: Doubt About Accuracy," Scientific American, 221 (December, 1969), 54.
- Truby, H. "Acoustico-cineradiographic Analysis Considerations," Acta Radiologica Supplementa, 182 (1959), 1-208.
- Turner, W. "Spectrogram Voice Identification," American Journal of Proof of Facts, 423 (1967), 431.
- Ungeheuer, C. "Two Dimensional Displays for Speaker Recognition," Study of Sounds, 12 (1966), 122-127.
- "U.S. v. Wright," Criminal Law Reporter, 1 (August 23, 1967), 2287.
- "Valuable Clues from Sounds Made Visible," Life, July 21, 1967, 56A-56B.
- "Visible Speech," Time, 46 (November 19, 1945), 50.
- Whevin, D. "Voice Detectives Go to Work on the Mystery Crash," Life, 56 (May 22, 1964), 46-46A.
- "Voiceprint Identification: Need for Further Experiments," San Diego Law Review, 6, 213.
- "Voiceprints Not That Reliable," Science Digest, 64 (January, 1970), 41.
- "Voiceprint Useful for Identification," Science News Letter, 81 (June 2, 1962), 343.
- "Voiceprints v. Criminals," Senior Scholastic, 82 (October 10, 1962), 19.
- "Voices Don't Lie: Research on Voice-prints," Newsweek, 59 (June 4, 1962), 62.

Voiers, W. "Perceptual Basis of Speaker Identity," Journal of the Acoustical Society of America, 36 (1964), 1065-1073.

Yanagihara, Naoaki. "Experimental Observation of the Noisy Quality of Harshness," Studia Phonologica, 3 (1964), 47-57.

Young, M. and Cambell, R. "Effects of Context on Talker Identification," Journal of the Acoustical Society of America, 42 (1967), 1250.

C. NEWSPAPERS

Borders, William. "Voiceprint as Evidence; Ruling Called First of Its Kind," New York Times, April 12, 1966, pp. 1-2.

Cohen, Jerry. "The Watts Voiceprint Case: A T.V. Boast that Trapped an Arsonist," Los Angeles Times West Magazine, March 26, 1970, pp. 12-19.

"Death Call of Airliner's Co-pilot," San Francisco Chronicle, June 4, 1964.

Gavzer, Bernard, "On Guard," State Journal (Lansing), September 25, 1970, p. A12.

Gilman, W. "Voice prints," New York Times Magazine, October 28, 1962, p. 74.

Kavanagh, Reginal, "'Voiceprint' Rules Out Di Gilio as Maker of Phone Threat," The Daily Home News (New Brunswick, New Jersey), March 7, 1970, p. 1.

"People v. Straehle and Rispole," New York Times, April 12, 1966, p. 1.

Prince, Walter, "Voiceprint: Tool or Toy?" Los Angeles Times West Magazine, March 26, 1967, p. 14.

"T.V. Remarks Lead to Watts Riot Conviction," Los Angeles Times West, December 10, 1966, p. 3.

"Two are Arrested for Bomb Scares at Pitney-Bowes," Stamford Advocate, July 2, 1964, p. 1.

"Voiceprints Nail Obscene Phone Caller," San Francisco Chronicle, May 27, 1966.

"Voiceprint System Studied," Army Times, March 18, 1970, p. 34.

"Voiceprint Unit Set Up," State Journal (Lansing), December 18, 1968.

"Voiceprint Used to Finger Rioter," Los Angeles Star-Ledger, December 19, 1966, p. 13.

D. UNPUBLISHED PAPERS

- Asp, Carl. "Voiceprint: Its Evolution, Application and Present Status," University of Tennessee.
- Beranek, Leo. "Acoustic Measurements." Office of Naval Research, Navy Department, Washington, D.C., 1956.
- Bolt, Richard ., and others. "On Speaker Identification by Speech Spectrograms: A Scientists View of its Reliability for Legal Purposes." Haskins Laboratories, New York. Report to the Technical Committee on Speech Communication of the Acoustical Society of America.
- Carbonell, J., and others. Final Report: Speaker Authentication Techniques. U.S. Army Electronics Laboratory, Fort Monmouth, New Jersey, May, 1965.
- Cederbaums, Juris. "Voiceprint Identification: A Scientific and Legal Dilemma." New York University School of Law, May 8, 1969.
- Clarke, F. R., Becker, R. W., and Nixon, J. C. "Characteristics that Determine Speaker Recognition." Report under contract to Stanford Research Institute, December, 1966.
- Denes, Peter. "The Speech Chain." Bell Telephone Series.
- Hecker, M. "Methods of Measuring Speaker Recognition." Stanford Research Institute, April, 1969.
- Kamine, Bernard. "The Voiceprint Technique of Speaker Identification: Its Validity and Admissibility in Court." March, 1968.
- Kersta, Lawrence, G. "Voiceprint In fallibility." A paper presented to the Acoustical Society of America, November 7, 1962.
- _____. "Voiceprint Classification for an Extended Population." A paper presented to the Acoustical Society of America, June 2, 1966.
- _____. "Automated Talker Identification by Quantized Spectrography." A paper presented to the Conference on Speech Communication and Processing, Air Force Cambridge Research Laboratories, Massachusetts, 1967.
- _____. "Voiceprint Classification." A paper presented to the Acoustical Society of America, February 14, 1967.
- _____. "Instruction and Application of Voiceprint Identification to Law Enforcement. A paper presented to the Acoustical Society of America, May, 1968.

- Kersta, L., and Colangelo, J. "The Spectrographic Speech Patterns of Identical Twins." Voiceprint Laboratories, Somerville, New Jersey, 1970.
- Kress, J. "Voiceprints and the Law." Columbia University School of Law, September, 1967.
- Nash, Ernest. Untitled paper. East Lansing, Michigan, 1969.
- Nash, Ernest. Personal interview. August 10, 1970.
- Presti, A. J. "A New Approach in Sound Spectrography." Voiceprint Laboratories, Somerville, New Jersey, February, 1970.
- "Procedures for Voiceprint Examinations." Voiceprint Laboratories, Somerville, New Jersey, 1969.
- Stapleton. Personal interview. September 11, 1970.
- "State v. Cary," Transcript of the New Jersey Supreme Court's opinion on the appeal by Paul Cary, A-155, September Term, 1966.
- Stewart, Don, and Bodfrey, John. "Voiceprint Identification." Laboratory for Experimental Phonology, School of Languages and Linguistics, Georgetown University, Washington, D.C.
- "Summary Review of Procedures for Speaker Recognition." Sensory Sciences Research Center, Stanford Research Institute, 1970.
- "Technical Aspects of Visible Speech." Monograph B-1415, Bell Telephone System, Murray Hill, New Jersey, November, 1957.
- Tosi, O. "An Evaluation of the Kersta Method of Voice Identification." Michigan State University.
- _____. "Lawrence G. Kersta, Method of Voiceprint Identification." Michigan State University, 1969.
- _____. "Speaker Identification Through Acoustic Spectrography." A paper presented to the XIV Congres De L' Association Internationale De Logopedie et Phoniatry, 1969.
- _____. Personal interview. August 17, 1970.
- Turner, Ralph, and Romig, C. Personal discussion. September 28, 1970.
- "Voice Identification Project." First Year Report of the Michigan State Police, East Lansing, Michigan, 1970.
- "Procedures for Voiceprint Examinations." Voiceprint Laboratories, Somerville, New Jersey, 1969.
- "Voiceprint Laboratories' Sound Spectrograph." Voiceprint Laboratories, Somerville, New Jersey, 1969.

Williams, C. "The Effects of Selected Factors on the Aural Identification of Speakers." A report to the Air Force Systems Command, Hanscom Field, Massachusetts, 1964.

E. LEGAL CASES

Aaron v. State, 271 Alabama 70, 122 So. 2d 360 (1960).

Beachem v. State, 144 Texas Criminal Review 272, 162 S.W. 2d 706 (1942).

Biggers v. Tennessee, 388 U.S. 909 87 S. Ct. 2132 (1967).

Boyd v. U.S., 292 L. Ed. 746 (1886).

Boyer v. State (Florida), 182 So. 2d 19 (1966).

Briethaupt v. Abram, 352 U.S. 432 (1957).

Cox v. State, Oklahoma Cr. 395 P. 2d 954 (1964).

Edison v. U.S. 272 F. 2d 684 (Tenth Cir., 1959).

Elimore v. Commonwealth, 282 Kentucky 443, 138 S.W. 2d 956 (1966).

Evans v. Commonwealth, 230 Kentucky 411, 19 S.W. 2d 1091 (1921).

Frye v. U.S., 293 F. 1013, 34 A.L.R. 145 (D.C. Cir., 1923).

Gilbert v. California, 388 U.S. 263, 87 S. Ct. 1951 (1967).

Griffin v. California, 380 U.S. 609, 85 S. Ct. 1229 (1965).

Johnson v. Commonwealth, 115 Pennsylvania 369, 9 A 78 (1887).

Katz v. U.S., 389 U.S. 347 (1967).

Lanford v. People, 159 Colorado 36, 409 P. 2d 829 (1966).

Lenoir v. State, 197 Maryland 495, 80 A 2d 3 (1951).

Miranda v. Arizona, 384 U.S. 436, 86 S. Ct. 1602 (1966).

McClard v. U.S., 386 F. 2d 495 (Eighth Cir., 1967).

Osborn v. U.S. 335 U.S. 323, 87 S. Ct. 429 (1966).

Palmer v. Peyton, 359 F. 2d 199 (Fourth Cir., 1966).

People v. Ellis, 65 California 2d 529, 55 California Reporter 385 (1966).

People v. Fisher, 340 Illinois 216, 172 N.E. 743 (1930).

People v. Jennings, 252 Illinois 534, 96 N.E. 1077 (1911).

People v. McKenna, 49 New Jersey Superior Court 71, 266 A 2d 757 (1967).

People v Morse, 325 Michigan 270, 38 N.W. 2d 322 (1949).

People v Sica, 112 California Appeals Court 2d 547.

People v Williams, 164 California Appeals Court 2d 858 (1958).

Robles v U.S., 297 F. 2d 401 (1960).

Rochin v California, 342 U.S. 165, 72 S. Ct. 205 (1952).

Schmerber v California, 384 U.S. 757 (1965).

State v Cary, 49 New Jersey 343, 351, 230 A. 2d 384 (1967).

State v Freeman, 195 Kansas 561 (1965).

State v McNamara, 104 N.W. 2d 568 (Iowa, 1960).

State v Ramirez, 76 New Mexico 72 (1966).

State v Roberts, 102 New Hampshire 414, 158 A 2d 458 (1960).

State v Taylor, 213 South Carolina 330, 49 S.E. 2d 289 (1948).

Stovall v Denno, 386 U.S. 293, 302, 87 S. Ct. 1967 (1967).

U.S. v Hiss, 88 F. Supp. 559 (1950).

U.S. v McKeever, 169 F. Supp. 426 (1958).

U.S. v Two Obscene Books, 99 F. Supp. 760 (1951).

U.S. v Wright, 17 U.S.C.M.A. 183, 37 C.M.R. 447 (1967).

Wade v. U.S., 388 U.S. 218 (1967).

Warden v Hayden, 387 U.S. 294 (1967).

APPENDIX A

SPECIFICATIONS OF THE VOICEPRINT

LABORATORIES' SOUND SPECTROGRAPH

Spectrogram Paper Size: 12 3/4 x 5 5/8 inches.

Input Tape: Standard 1/4 inch Mylar magnetic recording tape.

Analysis Segment: 2.4 seconds of taped signal is analyzed to 7,000 Hz
in 80 seconds.

Stylus Engagement: Automatic.

Stylus Pressure Adjustment: None. Stylus pressure is independent of
stylus wear.

Circuitry: All semiconductor electronics.

Filters: 45 Hz and 300 Hz analyzing filters.

Amplitude Display: Dynamic Range of 48 db's, quantized into eight
levels.

APPENDIX B

SPECIFICATIONS OF TAPE RECORDERS

Wollensak T-2,000

Frequency Response: 40 - 8,000 Hz, at 3 3/4 i.p.s.

40 - 15,000 Hz at 7 1/2 i.p.s.

Signal to Noise Ratio: 46 db's.

Wow and Flutter: 0.3% at 7 1/2 i.p.s.

Uher 4,000-1

Frequency Response: 40 - 17,000 Hz at 3 3/4 i.p.s.

40 - 20,000 Hz at 7 1/2 i.p.s.

Signal to Noise Ratio: 55 db's.

Wow and Flutter: 0.15 at 7 1/2 i.p.s.

APPENDIX C

FREQUENCY RESPONSE OF MICROPHONES

Electro-Voice Cardioid Unidirectional Microphone, Model RE 15

Frequency Response: 60 - 15,000 Hz.

Electro-Voice Omnidirectional Microphone, Model 6358

Frequency Response: 80 - 13,000 Hz.

Uher Dynamic, Omnidirectional Microphone, Model 515

Frequency Response: 70 - 14,000 Hz.

NOV 20 1970

MICHIGAN STATE UNIV. LIBRARIES



31293101815110