July 23, 1970

SUMMARY REVIEW OF PROCEDURES FOR SPEAKER RECOGNITION

Prepared by Staff

Sensory Sciences Research Center
Stanford Research Institute
Menlo Park, California 94025

I INTRODUCTION

When a person speaks he produces a complex acoustic signal that contains various kinds of information. This signal serves primarily to convey a linguistic message. Listeners who are familiar with the language can transcribe or at least repeat what the speaker said. Besides conveying a message the speech signal also reflects some of the anatomy and physiology of the speaker. For example, listeners can often determine the speaker's sex, his approximate age, his emotional state, and whether or not he is suffering from an illness (such as the common cold). Of particular interest is the ability of listeners to distinguish among the speech characteristics of different speakers. This ability is the basis of one method of speaker recognition.

There are three general methods of speaker recognition. These are speaker recognition by listening, speaker recognition by comparison of spectrograms, and speaker recognition by machine. Each of these methods is described in greater detail in separate sections of this report. Speaker recognition by listening is, of course, the method used in everyday life. It has been studied for a longer period of time and appears to be more accurate and reliable than either of the other methods as they are now practiced. A possible limitation of this method is that it is entirely subjective. No matter how accurate and reliable listeners may be they are usually unable to describe the criteria upon which their decisions are based and thus they are unable to justify their conclusions in a court of law.

Speaker recognition by visual comparison of spectrograms is considered to be a more objective method. Spectrograms are visual displays of the speech signal. They exhibit graphic features that can be discussed in a fairly objective manner. But these features are still interpreted subjectively in arriving at an overall decision. For this reason there has been much interest in a third method, namely, speaker recognition by machine. Although machine decisions are inherently objective, they are, as of now, often less accurate for speaker recognition purposes than comparable human decisions. Current research efforts in speaker recognition by machine are specifically directed toward overcoming this limitation.

All methods of speaker recognition are based on the fact that a given word or phrase tends to be uttered differently by different speakers. There is much variability in the speech signal and some of this variability is undoubtedly related to particular speaker differences. The nature of speaker variability is discussed as background material to provide the reader with an understanding of principles of speaker recognition.

II INTERSPEAKER AND INTRASPEAKER VARIABILITY

It is well-known that the pronunciation of a given word or phrase tends to vary from speaker to speaker. Acoustical analyses of utterances of several speakers typically reveal many dissimilarities. This effect is referred to as interspeaker (between-speaker) variability. Interspeaker variability in the speech signal can be attributed in part to organic differences in the structure of the vocal mechanism and, in part, to learned differences in the use of the vocal mechanism during speech production. Organic differences may be determined by heridity, sex, and age. Learned differences may be related to regional, social, and cultural factors.

Not so well-known is the fact that a particular speaker rarely utters a given word twice in exactly the same way, even when the utterances are produced in succession. This is referred to as intraspeaker (within-speaker) variability. In generating an utterance a speaker strives to produce appropriate respiratory, laryngeal, and articulatory activity that will lead to understandable speech. But many details of the resulting waveform will change from utterance to utterance depending upon rate of speaking, mood of the speaker, emphasis given to various words, and many other variables.

The success of any method of speaker recognition depends on the degree to which interspeaker variability is greater than intraspeaker variability. Both forms of speaker variability are extremely difficult to quantify, because speaker variability is a reflection of many differences in speech production. It cannot be meaningfully expressed in terms of a single measure. The measurement of speaker variability requires an understanding of how specific differences in speech production are manifested in the speech signal. But such an understanding is not yet available.

III SPEAKER RECOGNITION BY LISTENING

Several kinds of tests have been devised to study different aspects of speaker recognition by listening. All tests employ the same basic procedure. Speakers drawn from a prescribed population are recorded, while reading selected speech material. The recordings are edited and presented to listeners, and the listeners carry out a recognition task. Each step in this procedure introduces variables that can influence the resulting performance. These variables include the size and homogeniety of the speaker group, the selection of speech materials, the size and training of the listener group, the mode of presentation of speech material, and the specific task assigned to the listeners. Each of these classes of variables is discussed in some detail by Hecker (1970).

The objective of most studies on speaker recognition by listening is, of course, to appraise the likelihood that a listener's judgment might be in error. In fact one of the first studies of this kind was motivated by a legal question of fallibility that arose in the Lindbergh case of 1935 (McGehee, 1937). Lindbergh claimed that he recognized the voice of the defendant as the voice of his son's kidnapper, heard almost three years earlier. Although Lindbergh's testimony was accepted by the court, the defense argued that such recognition was not entitled to much weight as evidence.

McGehee studied the reliability with which listeners can recognize unfamiliar voices. Groups of listeners participated in two experimental sessions that were separated in time, from one day to five months. During the first session they heard an unfamiliar speaker read a paragraph of text. During the second session they heard the same paragraph read successively by five speakers, including the speaker from the first session. The ability of the listeners to recognize the speaker whom they heard in the first session was investigated as a function of the time interval between the two sessions. The results, which are shown in Table I indicate that the reliability of recognition decreases rapidly as the time interval between sessions is extended beyond two weeks.

Table I

Percent Correct Recognition of Unfamiliar Male Speakers After Various Intervals of Time (After McGehee, 1937.)

Days			Weeks			Months		
1	2	3	1	2	3	1	3	5
83%	83%	81%	81%	69%	51%	57%	35%	13%

The effect of increasing the number of speakers heard during the first sessions was also investigated. When one of two speakers heard during the first session spoke again during a second session two days later, 77 percent of the listeners recognized his voice. When five speakers participated in the first sessions, only 46 percent of the listeners could recognize one of their voices two days later. Vocal disguise was also found to be effective in lowering recognition scores. In this experiment only one speaker was heard during the first session. He disguised his voice by changing its fundamental frequency. During the second session he used his normal voice. With a time interval of only one day, correct recognition was reduced by 13 percentage points.

These results are illustrative of many of the results reported in the scientific literature. They illustrate the important fact that the speech waveform carries information relevant for distinguishing among talkers. However, the ability of listeners to identify speakers by their voice alone falls far short of 100 percent reliability. The quest for a more reliable means of identifying speakers on the basis of their voices has led to the study of speaker recognition by visual comparison of spectrograms and speaker recognition by machine. These two approaches will be briefly described in the following sections.

IV SEPAKER RECOGNITION BY VISUAL COMPARISON OF SPECTROGRAMS

This method of speaker recognition makes use of an instrument that converts the speech signal into a visual display. The instrument is called a sound spectrograph, and the display it provides is a sound spectrogram (or Voiceprint, a trade name owned by Voiceprint Laboratories, Somerville, New Jersey). Spectrograms of different utterances of a given word or phrase are presented to a trained observer who attempts to determine whether some utterances were produced by a common speaker. Because the method has obvious applications in criminology, many studies have been concerned with its reliability as a means of positive identification. The sound spectrograph consists of four basic parts: (1) a magnetic recording device, (2) a variable electronic filter, (3) a paper-carrying drum that is coupled to the magnetic recording device, and (4) an electric stylus that marks the paper as the drum rotates. The magnetic recording device is used to record a short sample of speech. The duration of the speech sample corresponds to the time required for one revolution of the drum. Then the speech sample is played repeatedly in order to analyze its spectral contents. For each revolution of the drum, the variable electronic filter passes only a certain band of frequencies, and the energy in the frequency band activates the electric stylus so that a straight line of varying darkness is produced across the paper. The darkness of the line at any point on the paper indicates how much energy is present in the speech signal at the specified time within the given frequency band. As the drum revolves, the passband of the variable electronic filter moves to higher and higher frequencies, and the electric 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.

Figure 1 shows three spectrograms. Since the spectrograms portray different utterances of the same phrase, each spectral feature of one utterance has a grossly similar counterpart in another utterance. The variability in corresponding spectral features appears to be somewhat

FIG. 1 SOUND SPECTROGRAMS OF THREE UTTERANCES OF THE PHRASE "MACHINE RECOGNITION OF SPEECH" (After Young and Hecker, 1968.)

greater between the two speakers (interspeaker variability) than between the two utterances by the same speaker (intraspeaker variability).

The spectrogram provides a permanent visual record of a speech signal. Such records may be studied in detail, point for point comparisons may be made among spectrograms, and judgments of similarity may be expressed in quantitative terms. Thus, the spectrogram has obvious appeal in legal applications. It is likely that the full potential of the spectrogram as a tool for achieving speaker recognition has not yet been reached.

However, the sound spectrogram has inherent limitations for speaker recognition applications. The display was designed to show differences among words and phonemes. It was not a purpose of the design to reveal differences between talkers. Thus, no attempt was made to have the device extract parameters from the speech waveform that might optimize speaker recognition performance. Further, a basic characteristic of all spectrum analyzers is that their frequency resolution can be increased only at the expense of temporal resolution and vice versa. The capability of a particular instrument to resolve frequency differences and temporal events is determined primarily by the bandwidth of its analyzing bandpass filter. Although the sound spectrograph contains two bandpass filters with different bandwidths (45 Hz and 300 Hz), the choice of either filter represents a compromise. Those features that might eventually prove to be the most useful ones for differentiating among speakers are not necessarily revealed in either the narrow-band or the wideband spectrogram.

Because of the limited resolving power of the sound spectrograph, it is possible that spectrograms prepared from slightly different utterances of the same word cannot be differentiated by human observers. While the differences among the utterances would be evident in oscillographic recordings (which describe the utterances most completely) these differences may be obscured in the sound spectrogram. Therefore, when two spectrograms appear to be identical in all respects, it cannot be concluded that they necessarily represent the same speech signal. This limitation can be particularly severe in cases where the speech signals under analysis are distorted, or embedded in noise.

The general procedure used in experiments employing the spectrogram as a means of speaker recognition is as follows: speakers are recorded while reading selected words or phrases. Spectrograms are prepared from the recordings. Two or more spectrograms of different utterances of the same words or phrases are presented to trained observers, and the observers carry out a recognition task. As is the case with speaker recognition by listening, each step in this procedure introduces variables that can affect performance, that is, the ability of the observer to match correctly spectrograms that represent the same speaker. The most important variables are described in detail by Hecker (1970), and will not be discussed in this report.

The fallibility of the observer is a crucial issue in the legal use of this method of speaker recognition (Borders, 1966; Ladefoged and Vanderslice, 1967; McDade, 1968; Bolt et al, 1970). Although a machine (the sound spectrograph) is used to prepare spectrograms, the interpretation of spectrograms is an art rather than a science. When this fact is pointed out to members of a jury they may be unable to evaluate the reliability of this means of identification. In the first trial in which spectrograms were allowed as evidence, the jury could not reach an agreement as to how much weight this evidence should be given (McDade, 1968). The conviction of Edward Lee King was reversed by a Court of Appeals because "The 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." (Kennedy, 1968)

The use of the term Voiceprint, and the degree to which the analogy between Voiceprints and fingerprints has been emphasized (Kersta, 1962a, 1962b; Anon., 1965; McDade, 1968) are rather unfortunate. There is an important difference between spectrograms and fingerprints that is too seldom considered. The intraspeaker variability of the speech signal can be substantial. And this variability is, of course, demonstrated in spectrograms that represent a particular speaker. The variability exhibited by the whorls and ridges on a particular person's fingers is essentially zero (Ladefoged and Vanderslice, 1967; Bolt et al, 1970). Any difficulty in matching fingerprints is caused by the fact that fingerprints may be incomplete or smeared. As a means of identification, fingerprints must be regarded as being considerably more foolproof than the spectrograms.

Claims by Kersta and others of the reliability of the Voiceprint for achieving speaker recognition are based largely on the results of unpublished experiments, thus the scientific community cannot appraise the design of these experiments and the validity of the conclusions reached (Ladefoged and Vanderslice, 1967). The published results of one series of experiments (Kersta, 1962b) could not be duplicated by other investigators. Young and Campbell (1967), and also Stevens, Williams, Carbonell, and Woods (1968), obtained much higher error scores than those reported by Kersta (1962a, 1962b). Such disagreements make the publication of detailed descriptions of future experiments extremely desirable and necessary.

In the first experiments concerned with the question of Voiceprint, the observers were required to sort spectrograms into groups that represented different speakers (Kersta, 1962a, 1962b). Later experiments employed the multiple-choice identification test (Kersta, 1962c; Young and Campbell, 1967; Stevens, Williams, Carbonell, and Woods, 1968). There have been no reports of experiments dealing directly with the type of identification task commonly encountered in criminal investigations. Ladefoged and Vanderslice (1967) argued that the reliability of Voiceprint identification in practical cases cannot be predicted from the results of the published studies.

It has been claimed that spectrogram recognition performance is essentially unaffected by the loss of teeth, tonsils, or adenoids, the aging process, and attempts to disguise the voice, such as changing the fundamental frequency, whispering, mimicking another voice, and ventriloquism (Kersta, 1962c, Anon., 1965). However, in the absence of supporting experimental data, these claims cannot be considered established facts. Furthermore, when the speech signal is degraded, as it may well be when transmitted by a typical telephone system, many of the above-mentioned factors can be expected to reduce the reliability of this method.

According to Kersta (1962b), the probability that two speakers have similar enough vocal-tract dimensions and articulation patterns to produce indistinguishable spectrograms is extremely small. This belief, which appears to underlie many experiments, has not been formally translated into a hypothesis that can be tested with a finite population of speakers. There is evidence that two arbitrarily selected speakers can occasionally produce very similar spectrograms (Lagefoged and Vanderslice, 1967). This situation is illustrated in Fig. 2 for the word "you." Findings of this kind suggest that the range of one speaker's pronunciations of a given word (intraspeaker variability) may partially overlap the range of another speaker's pronunciations of the same word, and argue for the use of a large number of different words in making an identification. There is also evidence of considerable similarity among spectrograms representing different members of a family (Kersta, 1965a), and this suggests another source of observer fallibility.

Stevens, Williams, Carbonell, and Woods (1968) examined the ability of observers to distinguish between familiar and unfamiliar speakers in a 32-item identification-discrimination test. The observer was given eight reference spectrograms that represented eight "familiar" speakers. There were two experimental conditions; either four or 16 of the 32 test spectrograms represented "unfamiliar" speakers who were not represented by the reference spectrograms. The results of this study are shown in Table II. Most of the familiar speakers were recognized as such, and they were subsequently correctly identified. Many of the unfamiliar speakers, however, were erroneously recognized as familiar speakers, especially when they appeared as often as the familiar speakers. As a point of comparison, listening tests were conducted using the same speakers and the same test format. Spectrograms were not employed in these tests. These data are shown in Table III. A comparison of the two sets of data reveals that there were considerably more acceptances of unfamiliar speakers in the visual tests than in the oral tests. When only four of the 32 test items represented unfamiliar speakers, there were also more false rejections of familiar speakers in the visual tests. Thus, speaker recognition by listening was found, in this study, to be the more accurate

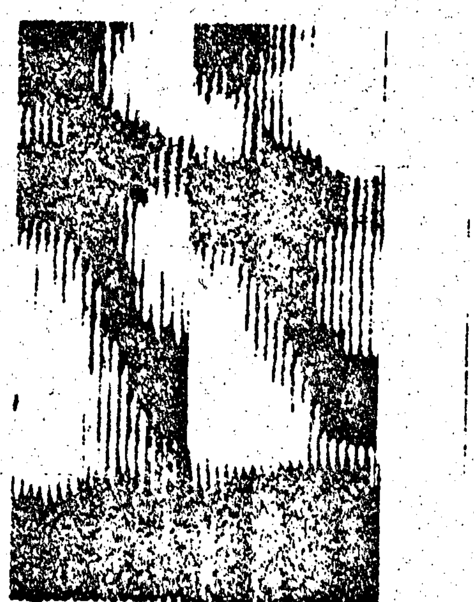

FIG. 2 SIMILAR SPECTROGRAMS OF THE WORD "YOU" UTTERED BY TWO ARBITRARILY SELECTED SPEAKERS (After Ladefoged and Vanderslice, 1967.)

Table II

Percent Correct Recognition of Familiar and Unfamiliar Male Speakers by Visual Comparison of Spectrograms (Data are shown for two experimental conditions. After Stevens, et al., 1968.)

4 of 32 Test Items by Unfamiliar Speakers		
	Recognized as	
Speaker	Familiar	Unfamiliar
Familiar	80	20
Unfamiliar	31	69
16 of 32 Test Items by Unfamiliar Speakers		
	Recognized as	
Speaker	Familiar	Unfamiliar
Familiar	90	10
Unfamiliar	47	53

Table III

Percent Correct Recognition of Familiar and Unfamiliar Male Speakers by Listening
(Data are shown for two experiment conditions. After Stevens et al., 1968.)

4 of 32 Test Items by Unfamiliar Speakers		
Speaker	Recognized as	
	Familiar	Unfamiliar
Familiar	88	12
Unfamiliar	6	94

16 of 32 Test Items by Unfamiliar Speakers		
Speaker	Recognized as	
	Familiar	Unfamiliar
Familiar	92	8
Unfamiliar	8	92

method. It must be pointed out that the observers employed by Stevens et al had very little training. One would expect better performance from highly-trained observers, but this study does demonstrate that speaker recognition by spectrogram matching is neither obvious nor easily achieved.

Data based upon carefully controlled experiment using well-trained observers will soon be available. In a program sponsored by the National Institute of Law Enforcement and Criminal Justice, U.S. Department of Justice, through the Michigan State Police, scientists at Michigan State University have been examining speaker recognition by visual comparison of spectrograms as a function of several variables including: quality of recordings, context of words used in the identification task, number of speakers in the comparison population, number of words used for identification purposes, and number of samples of each word (Tosi, 1970). These data, which should soon be published, will provide a good determination of the reliability of speaker recognition by the current technique of making visual comparisons of speech spectrograms.

The above discussion may be summarized as follows: In view of the use of the visual comparison of spectrograms for speaker identification as evidence in courts of law, the fallibility of the observer must be studied further (Bolt et al., 1970). Future experiments should be carefully designed so as to avoid possible artifacts in the results. A detailed description of the experimental procedure, accompanied by the obtained data, should be published or otherwise be made available to the scientific community. Claims should be clearly differentiated from proven facts, and statements establishing an analogy between Voiceprints and fingerprints should be avoided. Although the spectrographic method for speaker identification has obvious potential in various investigative and forensic applications, its reliability as a means of identification has not yet been established.

V SPEAKER RECOGNITION BY MACHINE

Two approaches have been used to study the feasibility of speaker recognition by machine. One approach is to have the machine generate and examine amplitude-frequency-time matrices of specific speech samples.

The other approach is to have the machine extract speaker-dependent parameters from the speech signal and subject them to a statistical analysis. Each approach has led to a number of recognition techniques.

In the first case, the utterances of specific speech samples are usually processed by a spectrum analyzer that consists of a bank of band-pass filters, rectifiers, and smoothing circuits. The outputs of the analyzer are periodically sampled, and the amplitudes are quantized for further processing by computer. Each utterance is represented in the computer by a data matrix. The rows of the matrix correspond to the frequency bands of the spectrum analyzer, the columns correspond to the temporal locations of the sample spectra, and each matrix cell contains the measured amplitude level. Such a matrix may be thought of as a "digital spectrogram." For each phrase, word, or phoneme used, several matrices representing different utterances by the same speaker are combined to form a single reference matrix for that speaker. A reference matrix is thus constructed for each speaker participating in a recognition experiment. The speaker to be recognized is represented by a test matrix. Depending on the type of recognition to be performed the test matrix is compared with all, or one of the reference matrices. The degree of similarity between the test matrix and each reference matrix is computed, and the results are used to arrive at a decision.

There are two basic recognition tasks, identification and discrimination. In the identification task, several reference matrices are used and it is assumed that the speaker represented by the test matrix is also represented by one of the reference matrices; thus, the reference matrix that is most similar to the test matrix is expected to identify the speaker represented by the test matrix. In the discrimination task only one reference matrix is used and the speaker represented by the test matrix may or may not be represented by this reference matrix. Decision rules are selected to specify when the test and reference matrices are similar enough to represent the same speaker. A summary description of six studies is presented in Table IV. For each experimental study this table gives the speech materials used, the configuration of the data matrix, the number of utterances included in the reference and test matrices, the recognition task, the number of speakers involved, and an

Table IV

Summary Description of Six Recognition Techniques Using Specific Cue Material

Study	Speech Material	Matrix Configuration				Utterances Incl.		Recogn. Task	Speakers	Perform. %
		Frequency Bands	Range kHz	Interval msec	Amplitude bits	Ref. Matrix	Test Matrix			
Pruzansky (1963	10 Words	17	0.2-7.0	10	10	3	1	Ident.	10	89
Pruzansky and Mathews (1964)	10 Words	17	0.1-10.0	10	10	3	1	Ident.	10	93
Ramishvili (1965)	10 Words	7	0.2-10.0	50	2	10	1	Ident.	20	92
Li et al. (1966)	3 Phrases*	15	0.3-4.0	20	12	10+	1	Discr.	20	90
Glenn and Kleiner (1968)	Conson. [n]	25	1.0-3.5	-	6	10	10	Ident.	30	93
Meeker (1967)	4 Vowels	19	0.2-8.0	40+	≁	20	3	Discr.	11	95

* Used only first 500 msec of each utterance.

≁ Used relative frequencies of occurrence of three spectral slopes.

overall measure of performance. Obtained percent correct scores range from 89 to 95 percent.

Techniques using statistical analyses of speech parameters involve two distinct processes: (1) the extraction from the speech signal of paramaters thought to be useful for differentiating among speakers, and (2) the application of decision rules to combinations of parameter values that represent particular speech samples.

Questions regarding the most appropriate speech parameters have generally not been resolved as well as have questions regarding optimal decision rules. Various kinds of parameters have been examined, using both waveform analyses and spectroanalyses of the speech signal. Studies conducted by Clarke and Becker (1969), Hargraves and Starkweather (1963), Smith (1962), Ramishvili (1966), Edie and Sebestyen (1962), and Floyd (1964) have considered many speech parameters and several decision techniques. In general, results have been promising but is clear that much work remains to be done before automatic recognition techniques attain high reliability.

VI FUTURE DEVELOPMENTS IN SPEAKER RECOGNITION

The previous material describes in general terms the current status of speaker recognition by listeners, by visual examination of spectrograms, and by machine. Here we will comment briefly on the potential of each of these methods.

(A) Speaker Recognition by Listeners

There is little likelihood that much can be done, or should be done, to improve the average individual's ability to recognize speakers by voice. Identification based on the average individual's recognition of voice will undoubtedly remain unreliable although in some cases it may be admitted as evidence. Trained linguists, on the other hand, are reported to be very good at recognizing various dialects and the geographical region of origin of speakers. They are sometimes employed in the investigation phase of law enforcement and have been used as expert witnesses in legal proceedings. It is very possible that some linguists are far superior to the untrained individual in achieving reliable speaker recognition. However, we know of no studies that have directly investigated this possibility, nor do we know of any plans to do so. Thus, it would appear that the potential of speaker recognition by listeners is quite limited.

(B) Speaker Recognition by Visual Examination of Spectrograms

It is unlikely that this method has achieved its full potential. There has been too little systematic study of spectrogram features to determine optimal procedures for discriminating among talkers. While the current performance of analyzing machines can undoubtedly be improved upon, the fact remains that the spectrograph was not designed to emphasize features useful for distinguishing among talkers and it discards much information that may be of value for this purpose. Whereas the speech spectrograph should prove to be an increasingly valuable tool for investigative purposes it is unlikely that it will ever, under all circumstances, permit positive identification by voice.

(C) Speaker Recognition by Machine

This method of speaker recognition should prove to be the most promising. Computers are now capable of performing fast and accurate analyses of speech waveforms. Various parameters may be abstracted from the speech waveform and analyzed to determine those features most useful for distinguishing among talkers. Freedom to choose these optimal parameters should enable machine performance to exceed that of listeners or of trained observers using spectrograms as these two latter methods suffer from strict and arbitrary limitations upon processing equipment. While it is not scientifically obvious that absolutely positive identification by voice alone will ever be achieved by any method, speaker recognition by machine has the best chance of attaining this goal. To achieve improved or perfect performance the relevant speech parameters must be properly identified and incorporated into the analysis and decision processes of the machine.

REFERENCES

Anonym, "Voice Print Identification." Criminalistics (W. W. Turner, ed.), Aqueduct Books, Rochester (1965).

Bolt, R. H., F. S. Cooper, E. E. David, Jr., P. B. Denes, J. M. Pickett, and K. N. Stevens, "Speaker Identification by Speech Spectrograms: A Scientists' View of its Reliability for Legal Purposes." J. Acoust. Soc. Am. 47, 597-612 (1970).

Borders, W., "Voiceprint Allowed as Evidence; Ruling Called First of Its Kind." The New York Times, Tues., April 12 (1966).

Clarke, F. R., and R. W. Becker, "Comparison of Techniques for Discriminating Among Talkers." J. Speech and Hearing Res. 12, 747-761 (1969).

Edie, J., and G. S. Sebestyen, "Voice Identification General Criteria." Rept. RADC-TDR-62-278, Rome Air Development Center, Air Force Systems Command, Griffiss AFB (May 1962).

Floyd, W., "Voice Identification Techniques." Rept. RADC-TDR-64-312, Rome Air Development Center, Research and Technology Division, Air Force Systems Command, Griffiss AFB (Sept. 1964).

Glenn, J. W., and N. Kleiner, "Speaker Identification Based on Nasal Phonation." J. Acoust. Soc. Am. 43, 368-372 (1968).

Hargreaves, W. A., and J. A. Starkweather, "Recognition of Speaker Identity." Language and Speech 6, 63-67 (1963).

Hecker, M. H. L., "Speaker Recognition." Monograph A.S.H.A., (1970) (In press)

Kennedy, H., "Appeals Court Reverses State's First 'Voiceprint' Conviction." Los Angeles Times, Fri., Oct. 11 (1968).

Kersta, L. G., "Voiceprint Identification." J. Acoust Soc. Am. 34, 725 (A) (1962a).

Kersta, L. G., "Voiceprint Identification." Nature 196, No. 4861, 1253-1257 (1962b).

Kersta, L. G., "Voiceprint-Identification Infallibility." J. Acoust. Soc. Am. 34, 1978 (A) (1962c).

Ladefoged, P., and R. Vanderslice, "The Voiceprint Mystique." Working Papers in Phonetics 7, University of California, Los Angeles (Nov. 1967).

Li, K. P., J. E. Dammann, and W. D. Chapman, "Experimental Studies in Speaker Verification, Using an Adaptive System." J. Acoust. Soc. Am. 40, 966-978 (1966).

McDade, T., "The Voiceprint." The Criminologist, No. 7, 52-60 (Feb. 1968).

McGehee, F., "The Reliability of the Identification of the Human Voice." J. Gen. Psychol. 17, 249-271 (1937).

McGehee, F., "An Experimental Study in Voice Recognition." J. Gen. Psychol. 31, 53-65 (1944).

Meeker, W. F., "Speaker Authentication Techniques." Tech. Rept. ECOM-02526-F, U.S. Army Electronics Command, Ft. Monmouth, N.J. (Dec. 1967).

Pruzansky, S., "Pattern Matching Procedure for Automatic Talker Recognition." J. Acoust. Soc. Am. 35, 354-358 (1963).

Pruzansky, S., and M. V. Mathews, "Talker-Recognition Procedure Based on Analysis of Variance." J. Acoust. Soc. Am. 36, 2041-2047 (1964).

Ramishvili, G. S., "Automatic Recognition of Speaking Persons." Rept. FTD-TT-65-1079, Foreign Technology Division, Air Force Systems Command, Wright-Patterson AFB (Dec. 1965).

Ramishvili, G. S., "Automatic Voice Recognition." Engineering Cybernetics, No. 5, 84-90 (Sept.-Oct. 1966).

Smith, J. E. K., "Decision-Theoretic Speaker Recognizer." J. Acoust. Soc. Am. 34, 1988 (A) (1962).

Stevens, K. N., C. E. Williams, J. R. Carbonell, and B. Woods, "Speaker Authentication and Identification: A Comparison of Spectrographic and Auditory Presentations of Speech Material." J. Acoust. Soc. Am. 44, 1596-1607 (1968).

Tosi, O., Personal communication (1970).

Young, J. R., and M. H. L. Hecker, "Some Observations on the Problem of Machine Recognition of Speech." Proc. 1968 National Electronics Conf., Chicago (Dec. 1968).

Young, M. A., and R. A. Campbell, "Effects of Context on Talker Identification." J. Acoust. Soc. Am. 42, 1250-1254 (1967).

McDade, T., "The Voiceprint." The Criminologist, No. 7, 52-60 (Feb. 1968).

McGehee, F., "The Reliability of the Identification of the Human Voice." J. Gen. Psychol. 17, 249-271 (1937).

McGehee, F., "An Experimental Study in Voice Recognition." J. Gen. Psychol. 31, 53-65 (1944).

Meeker, W. F., "Speaker Authentication Techniques." Tech. Rept. ECOM-02526-F, U.S. Army Electronics Command, Ft. Monmouth, N.J. (Dec. 1967).

Pruzansky, S., "Pattern Matching Procedure for Automatic Talker Recognition." J. Acoust. Soc. Am. 35, 354-358 (1963).

Pruzansky, S., and M. V. Mathews, "Talker-Recognition Procedure Based on Analysis of Variance." J. Acoust. Soc. Am. 36, 2041-2047 (1964).

Ramishvili, G. S., "Automatic Recognition of Speaking Persons." Rept. FTD-TT-65-1079, Foreign Technology Division, Air Force Systems Command, Wright-Patterson AFB (Dec. 1965).

Ramishvili, G. S., "Automatic Voice Recognition." Engineering Cybernetics, No. 5, 84-90 (Sept.-Oct. 1966).

Smith, J. E. K., "Decision-Theoretic Speaker Recognizer." J. Acoust. Soc. Am. 34, 1988 (A) (1962).

Stevens, K. N., C. E. Williams, J. R. Carbonell, and B. Woods, "Speaker Authentication and Identification: A Comparison of Spectrographic and Auditory Presentations of Speech Material." J. Acoust. Soc. Am. 44, 1596-1607 (1968).

Tosi, O., Personal communication (1970).

Young, J. R., and M. H. L. Hecker, "Some Observations on the Problem of Machine Recognition of Speech." Proc. 1968 National Electronics Conf., Chicago (Dec. 1968).

Young, M. A., and R. A. Campbell, "Effects of Context on Talker Identification." J. Acoust. Soc. Am. 42, 1250-1254 (1967).