# Speaker Identification (PART 2)

## Results of the National Academy of Sciences' Study

By BRUCE E. KOENIG
*Special Agent*
*Technical Services Division*
*Federal Bureau of Investigation*
*Washington, D.C.*

The first part of this article dealt with the three general categories of speaker identification—listening, machine analysis, and aural-visual comparison. In view of the recent increase in the use of voice identification technology and the problems that have ensued, a study was conducted by the National Academy of Sciences (NAS). Part II of this article reports on the committee's findings and recommendations.

During the past 15-20 years, the Federal Bureau of Investigation has witnessed the expanded use of voice identification technology, conflicting scientific opinions, court rulings for and against evidence based on sound spectrograms, and the establishment of an organization of voiceprint examiners. Therefore, in March 1976, the FBI requested that the NAS undertake a study on the practice and use of spectrographic or voiceprint identification, its reliability, and its use as evidence in court.

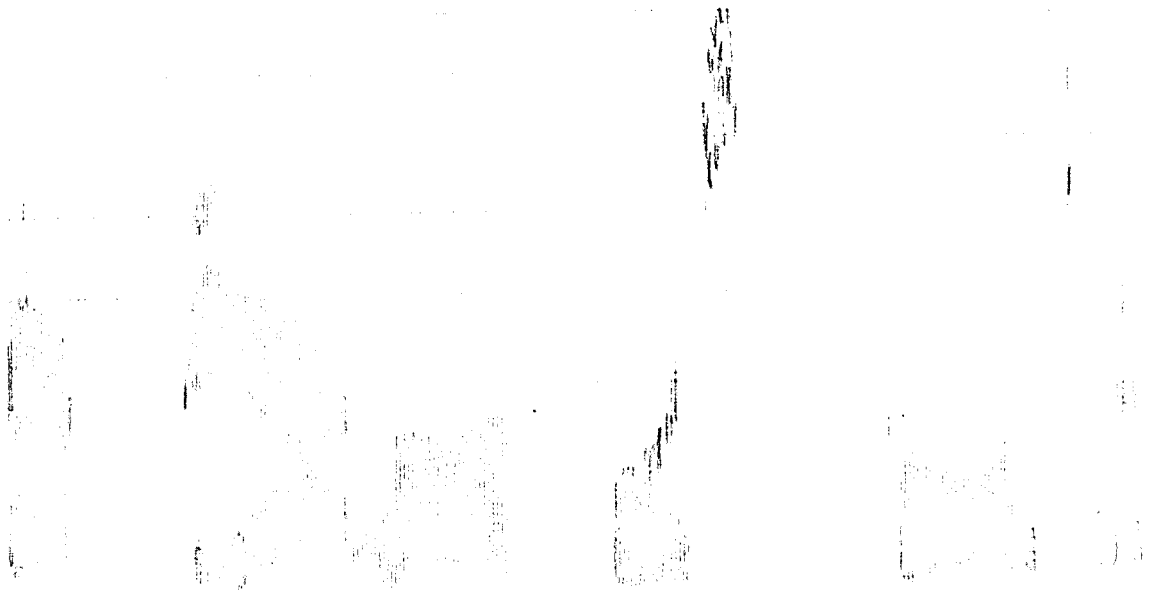
In July 1976, the National Research Council of NAS appointed the Committee on Evaluation of Sound Spectrograms which included eight independent experts representing both the scientific and legal communities. The committee included experts on acoustics, speech science, speech pathology, electronics, electrical engineering, audio recording systems, and criminal law and laws of evidence.

In February 1979, a detailed report of the committee's findings on voice identification, entitled "On the Theory and Practice of Voice Identification," was published. The committee throughout the study uses the term "voicegrams," which is synonymous with spectrograms or voiceprints. The committee did not address the issue of voiceprint admissibility in courts of law, which it considered to be a responsibility of the judicial and legislative bodies.

### Committee Findings

The findings of the study discuss, in part, three general areas:

1) Some information on the identity of an individual is obtainable through listening and by looking at voicegrams of that person's speech.



2) Voicegrams are fundamentally different from fingerprints in that fingerprints are unchanging for an individual, whereas the same word changes acoustically, at least slightly, every time it is spoken by a particular person.

3) "The degree of accuracy, and the corresponding error rates, of aural-visual voice identification vary widely from case to case, depending upon several conditions including the properties of the voices involved, the conditions under which the voice samples were made, the characteristics of the equipment used, the skill of the examiner making the judgments, and the examiner's knowledge about the case. Estimates of error rates now available pertain to only a few of the many combinations of conditions encountered in real-life situations. These estimates do not constitute a generally adequate basis for a judicial or legislative body to use in making judgments concerning the reliability and acceptability of aural-visual voice identification in forensic applications." [7]

## Conclusions of the Committee

The committee listed the three following conclusions in the areas of practice, research, and forensic use:

1) ". . . some improvement in the practice of aural-visual voice identification could be achieved in the near term by applying knowledge and techniques that are available now. . . .

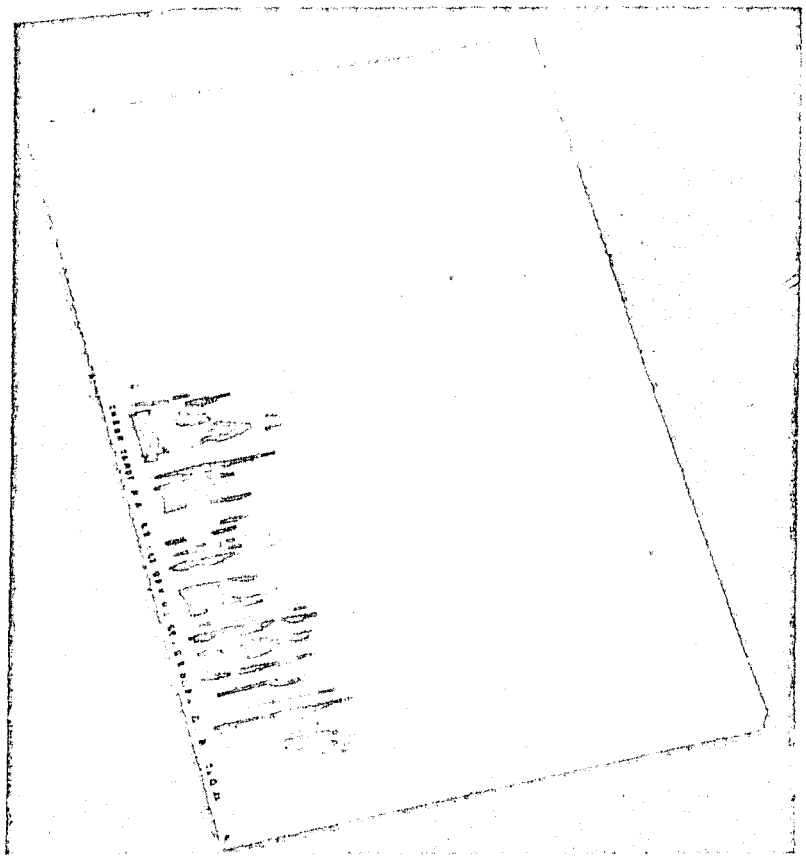
2) ". . . the full development of voice identification by both aural-visual and automated methods can be attained only through a longer-term program of research and development. . . .

3) "The decision about whether to use the aural-visual method of voice identification for forensic purposes depends on the answers to several subsidiary questions. First, it is necessary to have some measure of the error rate associated with the technique. . . . but objectively justified error rates are virtually impossible to determine for most of the forensic experiences reported to date." [8] Second, it is necessary to decide whether, in principle, the error rate is acceptably low for use in the particular case, which "is a value question and not a question of scientific or technical fact." [9] Third, it is necessary to decide whether, in practice, the nature of the error rate and the possible sources of error can be explained adequately to the lay fact finder, whether judge or jury, who will decide the case. "Therefore any presentation of voicegram evidence should be accompanied by a clear and thorough explanation of the limits of present knowledge about the accuracy of the technique. Such an explanation under present circumstances may be impossible to achieve or at least unwieldy, or it may be very costly." [10]

## Committee Recommendations

The committee made the following four recommendations regarding the use and practice of voice identification:

1) "We recommend that a mechanism be established to stimulate, guide, and coordinate a broad national program of scientific research on the processes of speech generation, transmission, and analysis as they pertain to the practice of voice identification. . . .



2) "We recommend that a national mechanism be established to develop objective standards and methods for testing the performance of voice identification examiners and to certify their competence as examiners. An existing organization, the International Association of Voice Identification (IAVI), was established to perform some of these functions. However, the Committee believes that [the] IAVI as presently constituted does not possess the broad base of representation usually considered appropriate and perhaps essential for a national certifying board. . . .

3) "We recommend that practitioners of aural-visual voice identification make full use of certain available knowledge and techniques that could improve the voice identification method. . . .

4) "We recommend that if evidence on voice identification is admitted in court—and we take no position on admissibility—then the inherent limitations in the method and in the performance of examiners should be explained to the fact finder, whether the judge or the jury, in order to protect against overvaluation of such evidence. . . . the testimony should explain that up to the present time, error rates for voice identification have been measured for only a limited number of experimental conditions. All the scientific results and forensic experiences to date, taken together, do not constitute an adequate objective basis for determining the error rates to be expected for voice identification testimony given in forensic cases generally. Error rates reported in specific cases cannot be much more than informed guesses based on practical experience combined with fragmentary results from scientific experiments. . . . Those limitations bear directly upon the problem of overvaluation of technical evidence." [11]

### Committee Summary

The summary of the NAS study states, in part, the following:

"The practice of voice identification rests on the assumption that intraspeaker variability is less than or different from interspeaker variability. However, at present the assumption is not adequately supported by scientific theory and data. Viewpoints about probable errors in identification decisions at present result mainly from various professional judgments and fragmentary experimental results rather than from objective data representative of results in forensic applications.

"The Committee concludes that the technical uncertainties concerning the present practice of voice identification are so great as to require that forensic applications be approached with great caution. The Committee takes no position for or against the forensic use of the aural-visual method of voice identification, but recommends that if it is used in testimony, then the limitations of the method should be clearly and thoroughly explained to the fact finder, whether judge or jury." [12]

**FBI**

*Copies of the NAS study are available for $7.00 prepaid through the Office of Publications, National Academy of Sciences, 2101 Constitution Ave., Northwest, Washington, D.C. 20418.*

*The FBI conducts voice identification examinations for Federal, State, and local law enforcement authorities for investigative guidance only and will not provide expert testimony.*

**Footnotes**

[7] *On the Theory and Practice of Voice Identification* (Washington, D.C.: National Academy of Sciences, 1979), p. 60.
[8] *Ibid.*, pp. 60-62.
[9] *Ibid.*, p. 62.
[10] *Ibid.*, p. 63.
[11] *Ibid.*, pp. 63-69.
[12] *Ibid.*, p. 2.

END