J Forensic Sci, Jul. 1973, Vol. 18, No. 3

J. W. Gocke,[1] *Captain, U.S. Army Judge Advocate General Corps and W. A. Oleniewski,*[1] *Captain, U.S. Army Judge Advocate General Corps*

Voiceprint Identification in the Courtroom

In a recent article appearing in the Journal of Forensic Sciences, J. J. Hennessey and C. H. Romig [*1*] presented a review of the experiments involving identification by voiceprint spectrograms, concluding that the various experiments in the field were conflicting, neither denying nor confirming the validity of voiceprint identification. This disagreement among scientists has had a tremendous affect on the admissibility of voiceprint identification as evidence in judicial proceedings. With one exception, until the end of 1970 evidence of voiceprint identification has been held inadmissible [*2*]. Since that time, the results of an extensive experiment on voice identification conducted by Dr. Oscar Tosi, Professor of Audiology and Speech Sciences, Michigan State University, have been released and greatly strengthened the argument for the admissibility of voiceprint identification as competent evidence [*3*].

In its simplest terms, a voiceprint spectrogram is a translation of sound into a pictorial representation by electronic means. When the sound in question is the spoken word, the pictorial representation is called a voiceprint. The primary work relating to the use of voiceprint spectrograms as a means of identification was done by Lawrence G. Kersta. From his experiments Mr. Kersta concluded that each person's voice is unique, and this uniqueness is imparted to and revealed in the pattern created when the voice is translated by the spectrograph into a voiceprint. He professed that, given the benefit of a five-word comparison, he could be 99.65 percent accurate in voiceprint identification [*2*].

Since a voiceprint spectrogram is proffered as evidence to show the similarity between an unknown voiceprint and that of a known person, the main dilemma facing the court is the reliability of voiceprints as competent scientific evidence [*4*]. While voiceprints fall within the purview of scientific tests, their admissibility will hinge largely upon the acceptance of their reliability by the scientific community [*5*]. The reliability of Kersta's experiments failed to convince the scientific community as a whole. The two main objections to his experiments were that he used a heterogeneous sampling of unknown voices; that is, the spectrograms used represented speakers with different accents, of different ages and backgrounds, making it easier to differentiate btween speakers; also his experiments were conducted using only closed testing groups, that is, the spectrograms of the unknown voices were always included in the group of the known voices being used. As a result, with closed groups, an examiner would merely have to find the sample in the known group most

The opinions or assertions contained here are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

[1] Formerly Fellows in Forensic Medicine, Armed Forces Institute of Pathology (AFIP), Washington, D.C. 20305. Captain Gocke is presently Special Assistant for Legal Affairs, AFIP, and Captain Oleniewski is Assistant Chief, General Claims Division, U.S. Army Claims Service, Ft. Meade, Md.

closely matching the spectrogram from the unknown group in order to make an identification [4]. These objections were expressed in two criminal cases in which voiceprint identifications were in issue.

State v. Cary [6] involved an interlocutory appeal in a murder trial from a pretrial order for the defendant to submit to a recording of his voice. The recording was to be eventually converted into a voiceprint to be compared with the unidentified voiceprint of the murder suspect. While the court affirmed the order to submit to a voice recording, it left unanswered the admissibility of the voiceprint as evidence until the demonstration of its reliability was presented; the court declared [6],

> . . . the prosecutor must satisfy the trial judge that identification by voiceprint technique and equipment has a sufficient scientific basis to produce uniform and reasonably reliable results and will contribute materially to the ascertainment of truth.

The trial court subsequently refused to admit the voiceprint into evidence, concluding that the voiceprint technique would not "as of this time" [7] produce reasonably reliable results.

The California case of *People v. King* [8] also denied admission of voiceprints as evidence.

In *King*, the defendant was convicted of arson arising out of the 1965 riots in Watts. His identity had been revealed by a comparison of voiceprints, one of which was made from the voice of an unidentified individual in a CBS television interview who had made some self-incriminating remarks regarding a number of arson-created fires; the second voiceprint was taken from the defendant once he had been considered to be a suspect. The government witness, Mr. Lawrence G. Kersta, identified the defendant's voice as being the same voice that made the incriminating statements in the CBS interview. On appeal, Mr. Kersta's opinion regarding the voiceprint was declared to have been inadmissible evidence, and the conviction was reversed. Essentially, the reason that prompted the appellate court into excluding Kersta's identification was that it was not convinced that Kersta's experiments in the field of voiceprints had crossed the line between the experimental and demonstrable stages or that the use of spectrograms had gained acceptance by competent scientific authorities.

In both of these decisions, the courts relied heavily upon the testimony of scholars in the field of acoustics. At this point in the development of voiceprint identification, many scientists felt that too much weight was being given to the spectrograph experiments of one worker, namely Mr. Lawrence G. Kersta. Many criticized his claim of almost absolute identification by the use of voiceprints as being unscientific.

The Acoustical Society of America's Technical Committee on Speech and Communications made a study of the legal implications of speaker identification by the Kersta method and concluded that the available results were inadequate to establish the reliability of voice identification by spectrograms. The Committee called for more controlled experiments to include the matching of voices from a group of spectrograms in which the unknown voice may or may not be present among the known voices [9]. Despite this scientific criticism, the United States Court of Military Appeals, in *United States v. Wright* [2], permitted the use of voiceprint spectrograms as evidence.

Wright was charged with five specifications of communicating obscene language to a female and five specifications of communicating a threat. The court found him not guilty of one of the obscene language specifications and one of the threat specifications but convicted him of the remaining eight specifications. During the trial Lawrence G. Kersta was called by the prosecution to testify that it was his opinion that an obscene and threatening telephone call made to the victim, which had been recorded on magnetic tape, was

made by the accused. Mr. Kersta identified the accused's voice on this tape by comparing it with a known exemplar of the accused's voice, which had also been preserved on magnetic tape.

After an exhaustive discussion of the development and use of voice identification by spectrograph was undertaken, the court held that the voiceprints were admissible. The court was convinced that Kersta's work was sufficiently explained by him to be based on sound scientific knowledge and had crossed from the experimental and demonstrable states into the "twilight zone" of evidential acceptance. Scientific disagreement would affect the weight to be given the evidence, but not its admissibility as evidence.

Since the decisions in the three cases mentioned above, Dr. Tosi reported the findings of his extensive study, which remedied the major defects of Kersta's work. Dr. Tosi used homogeneous samplings of unknown voices in both open and closed experiments. From more than 25 000 students at Michigan State University, he selected 250 students who spoke nonaccented English, had no noticeable speech defects, were all male undergraduates, and ranged in age from 19 to 24. The examiners were told that the spectrograms of the unknown voices might or might not be among the spectrograms of the known speakers. In that way, if the spectrogram of an unknown speaker did not match one of the known speaker's spectrograms, the examiner would not make an attempt either to identify or to pair up the two most closely resembling voiceprints. This open testing group would more closely parallel an actual situation in which law-enforcement officials would attempt to make a voice identification without knowing whether a suspect's voice would match that of an unknown speaker. A total of 34 996 experimental trials of identification were performed by 29 trained examiners. Each task involved up to 40 known voices, in various conditions: closed and open trials, contemporary and noncontemporary spectrograms, 9 or 6 clue words spoken in isolation and in fixed context and in random context. The examiners were forced to reach a positive decision (identification or elimination) in each instance, taking an average time of 15 minutes. Their decisions were based solely on inspection of spectrograms; listening to the voices was discarded from this experiment. Experimental tasks of this experiment, correlated with forensic models yielded an error of approximately 6 percent of false identifications and approximately 12 percent of false eliminations. The examiners judged approximately 60 percent of their wrong answers and 20 percent of their right answers as "uncertain." This suggests that if the examiners had been able to express no opinion when in doubt, on 74 percent out of the total number of tasks would have had a positive answer with approximately a 2 percent of errors of false identification and a 5 percent error of false eliminations [*3*]. Thus relying on these two years of experimentation and nearly 35 000 separate voice-identification trials, Dr. Tosi concluded that voice identification through spectrogram analysis has "a definite usefulness in the investigation of a crime" [*4*].

In light of the results of the experiment and field study, Dr. Tosi concluded that the voiceprint method could be employed as an identification tool if the following standards were maintained [*3*]:

1. Examinations must include both aural and visual comparisons.
2. The examiner must be a qualified professional, trained in phonetics and speech sciences. A two-year apprenticeship in field work should be required, along with academic training, to earn professional certification.
3. If he has the least doubt, the professional examiner must abstain from offering any positive conclusion. Since the voiceprint method of identification relies heavily on the expertise of the examiner, prudence should be the cardinal principle that guides the examiner's decision.

4. The examiner must be entitled to spend as much time and to use as many samples as he deems necessary to reach a conclusion.

Trimble v. Hedman was the first reported opinion reviewing the conclusions of Dr. Tosi's study, and ruled that voiceprints were admissible as evidence to corroborate voice identification by ear [*10*]. In that case, the St. Paul, Minnesota, Police Department received a telephone request for emergency assistance of a pregnant woman. When the police arrived at the scene, one of the police officers was ambushed and fatally injured by a bullet from a high-caliber rifle.

All emergency calls received by the St. Paul Police Department are recorded on tape, and that tape which contained the message which lured the police to the scene of the homicide was analyzed by a voiceprint spectrograph. Subsequently, the voiceprint from the anonymous call was compared to a number of other voiceprints taken from suspects of the killing, and the results indicated that the defendant *Trimble* had made the anonymous call.

Essentially, this was the same utilization made of a voiceprint spectrogram as in *Wright*. The Minnesota Supreme Court held that voiceprints can be used at least to corroborate voice identification by ear, if a proper foundation is first laid in establishing the expertise of the person preparing the spectrogram. The court also stated that voiceprints should be admissible for the purpose of impeachment and that information including a comparison of tapes and spectrograms could furnish probable cause for the arrest of a suspect. Though strong reliance was placed upon the credentials of the experts who testified, the *Trimble* court would not permit voiceprints to be the sole means of identifying a suspected criminal. Nonetheless, it did mention that the experiments conducted by Dr. Tosi were sufficient to show that voiceprint identification had reached such a scientific basis as to produce uniform and reasonably reliable results.

In a recent decision, *United States v. Raymond* [*4*], Judge Oliver Gasch was even more convinced of the reliability of voiceprints than the Minnesota Court.

Raymond was charged with shooting a policeman as he responded to an emergency telephone request for help. After the defendant was arrested, his voiceprint was compared to the voiceprint of the person making the false call, which had been routinely recorded along with all other incoming calls to the Washington Metropolitan Police Department. On the basis of comparison, voiceprint experts concluded that the two recordings were made by the same individual.

In his memorandum opinion, Judge Gasch stated [*4*]:

> It is on the basis of the extensive Tosi study, his testimony in open court, and the opinions expressed by other experts, that this court concludes spectrogram analysis is admissible evidence.

In support of its conclusion, the court made reference to the expertise of Lieutenant Ernest Nash, a voice identification technician with the Michigan State Police Department, who made the identification of Raymond, stating [*4*]:

> In an actual forensic situation, an experienced examiner like Lt. Nash will only make an identification when he feels a high degree of certainty. For example, out of some 1250 examinations performed by Nash in which spectrograms of an unknown speaker were compared to those of a known speaker, Nash made only about 180 positive identifications, eliminated positively about 450 and would not make a definite decision in the remaining 620 some odd comparisons. This is one of the significant factors which led Dr. Tosi to state that the possibility of Nash making a mistaken identification is 'negligible'.

This is the only opinion that places voiceprint identification in an equal status with other scientific identification such as handwriting analysis, neutron activation analysis, and ballistics tests.

Perhaps the prime indication that voiceprint identification has been accepted by the scientific community is the conversion of the major opponents of the system into its most outspoken proponents. Prior to completing his own work, Dr. Tosi and Dr. Peter Ladefoged, Professor of Phonetics at U.C.L.A., testified against the scientific reliability of voiceprint identification in those earlier cases that excluded its admission into evidence [*7,8*]. Since the promulgation of Dr. Tosi's study, both men have appeared in court and testified on behalf of voiceprint reliability [*4,10*].

In addition to *Wright*, *Raymond*, and *Trimble*, as of March 1972 voiceprints have been offered and accepted as legal evidence in several cases [*11*]. In each, the standards that Dr. Tosi deems absolutely necessary to maintain in order that voiceprints be acceptable as a means of identification were explained to the courts. After lengthy examinations and cross-examinations, the voiceprint evidence was accepted in all of these cases. In view of these decisions, the dilemma of scientific disagreement as to the validity of voiceprint identification no longer seems to be an obstacle to a more widespread use of this technique in the courtroom.

References

[*1*] Hennessey, J. J. and Romig, C. H., "A Review of the Experiments Involving Voiceprint Identification," *Journal of Forensic Sciences*, Vol. 16, April 1971, pp. 183–198.
[*2*] *United States v. Wright*, 37 CMR 447 (Court of Military Appeals 1967).
[*3*] Tosi, Oscar, "Experiments on Voice Identification," *Journal of the Acoustical Society of America*, Vol. 51, June 1972, pp. 2030–2043.
[*4*] *United States v. Raymond*, 337 F. Supp. 641, 643, and 644 (Dist. Ct., D.C. 1972).
[*5*] *Frye v. United States*, 293 F. 1013 (U.S. Ct. App. D.C. 1923).
[*6*] *State v. Cary*, 230 A. 2d 384 (N.J. 1967).
[*7*] *State v. Cary*, 239 A. 2d 680, 685 (N.J. 1968).
[*8*] *People v. King*, 266 Cal. App. 2d 437 (Cal. 1968).
[*9*] Bolt, Richard H., et al, "Speaker Identification by Speech Spectrograms: A Scientist's View of Its Reliability for Legal Purposes," *Journal of the Acoustical Society of America*, Vol. 47, Feb. 1970, pp. 597–612.
[*10*] *Trimble v. Hedman*, 192 N.W.2d 432 (Minn. 1971).
[*11*] Unreported decisions in Florida (*State v. Worley, State v. Alea*, 1971); Missouri (*State v. Crowe*, 1971); Illinois (*State v. Merholc*, 1971); and California (*State v. Coffey*, 1972).