1972

Evidence: Admissibility of Spectrographic Voice Identification

Minn. L. Rev. Editorial Board

Case Comment

Evidence: Admissibility of Spectrographic Voice Identification

I. THE TRIMBLE CASE

In response to an emergency telephone call, the St. Paul, Minnesota Police Department dispatched two officers to assist an expectant mother. The officers did not observe anyone in need at the address given, but while at the scene, one of them was mortally wounded by a shot fired from ambush. The St. Paul Police Department routinely records all emergency telephone calls. The tape of the luring call was sent to the Michigan State Police where an acoustic spectrogram was made. Subsequently, petitioner became a suspect. With prior judicial approval and the cooperation of the Welfare Department, the police tape recorded a telephone conversation with petitioner. This tape was also forwarded to the Michigan State Police where another acoustic spectrogram was made. By comparing the two spectrograms, an officer of the Michigan State Police identified petitioner as the person who had made the luring call to the police. With this identification serving as probable cause, a warrant was issued for petitioner's arrest. After arrest and later indictment for first degree murder, petitioner applied for a writ of habeas corpus claiming the warrant for her arrest was issued without probable cause. The district court dismissed the writ and the Minnesota Supreme Court affirmed, *holding* that the spectrographic voice identification justified the issuance of the warrant. In addition, the court went on to express an opinion that spectrograms ought to be admissible at trial to corroborate aural identification. *State ex rel. Trimble v. Hedman*, 192 N.W. 2d 432 (Minn. 1971).

The Minnesota court in this decision became the first nonmilitary court of final appeal to approve the use of spectrograms for any purpose.[1] In evaluating the propriety of the approval of

1. Because the rule adopted in this case was expressed in dictum, the issue might have arisen again on appeal since the spectrographic evidence was used in the trial. Minneapolis Tribune, March 9, 1972, § B, at 1, col. 3. However, the defendant was acquitted. Minneapolis Tribune, March 17, 1972, § A, at 1, col. 7.

this identification technique at trial, this comment will review the prior cases, describe the spectrogram voice identification technique and the available experimental data, and, finally, comment on the extent to which the technique satisfies the standard for admission of scientific evidence.

II. THE PRIOR CASES

Although the use of spectrographic voice identifications evidence is apparently growing,[2] there are in addition to *Trimble* only three reported cases dealing with the issue.[3] In a military case, which admitted such evidence, there is no real discussion of either the standard for admission of scientific evidence or the theory and the then existing experimental data pertinent to this identification technique. The other two cases rejected the evidence and contain only brief descriptions of the theory, process and purported experimental support.[4]

The United States Court of Military Appeals in *United*

---

2. The technique involved in this comment is only mentioned in a footnote in 3 J. WIGMORE, EVIDENCE § 795a, at 251 n.10 (3rd ed. Chadbourne rev. 1970).

Besides the materials cited in WIGMORE, other legal writing on the subject includes: Kamine, *The Voiceprint Technique: Its Structure and Reliability*, 6 SAN DIEGO L. REV. 213 (1969). This article is an excellent explanation of the technique and review of the state of knowledge in 1969. The author concludes that at the time there was insufficient acceptance to warrant admitting spectarographic evidence but he points to the MSU Experiment, then underway (*see* text accompanying note 31 *infra*), as a potential watershed event. *See also* 19 AM. JUR. PROOF OF FACT 423. While this material is correct in its technical portions, it represents considerable overstatement as to the accuracy of the process.

The spectrographic voice identification technique has apparently been sanctioned to some degree by the proposed amendment to FED. R. CRIM. P., Rule 41.1. *See* Note, *Proposed Federal Rule of Criminal Procedure 41.1*, 56 MINN. L. REV. 667 (1972).

Much of the literature uses the term "voiceprint" with the connotation that the technique is similar to fingerprints, but this does not seem to be the case. *See* M. HECKLER, SPEAKER RECOGNITION: AN INTERPRETIVE SURVEY OF THE LITERATURE, 168-69 (Am. Speech and Hearing Ass'n Monographs No. 16, 1971); O. TOSI, H. OYER, W. LASHBROOK, C. PEDREY, J. NICHOL & E. NASH, AN EXPERIMENT ON VOICE IDENTIFICATION 3 (Mich. State Univ., Excerpts from the Report SHSLR 171, 1971) [hereinafter cited as MSU EXPERIMENT]. Accordingly, this comment describes the technique as spectrographic voice identification.

3. United States v. Wright, 17 U.S.C.M.A. 183, 37 C.M.R. 447 (1967); State v. Cary, 99 N.J. Super. 323, 239 A.2d 680 (Sup. Ct. 1968), *aff'd per curiam*, 56 N.J. 16, 264 A.2d 209 (1970); People v. King, 266 Cal. App. 2d 437, 72 Cal. Rptr. 478 (1968).

4. This experiment is discussed in note 28 *infra*.

States v. Wright,[5] a case involving obscene telephone calls, held spectrographic identification evidence admissible. The court's opinion is confusing since it is based upon the qualification of an expert rather than the accuracy and acceptance of the expert's scientific technique.[6] The court admitted the evidence because the tapes used to make the spectrograms themselves were available to the jury and "[s]ince voice identification by ear is fully acceptable in the courts, the court members (i.e., the jury) could thus determine for themselves the margin of error, if any, in [the expert's] opinion."[7] The approach established by this language seems to be one of "admit it and let the jury decide." The dissenting judge in *Wright* stated that the scientific acceptability of spectrographic identification was in issue, not the qualification of an expert, and that the evidence in the case did not show acceptance by the scientific community.

State v. Cary[8] was a murder case in which the police had a tape of a telephone call received by them regarding the crime. The defendant appealed from a pre-trial order that he submit to a blood test and a recording of his voice for the purpose of spectrographic identification. The New Jersey Supreme Court remanded the case to the trial court to determine whether spectrographic identification would be admissible. After setting forth the theory underlying the technique and the experiment[9] and noting the contradictory testimony on the accuracy of the technique, the trial court held such evidence inadmissible since its accuracy was not generally accepted by the scientific community.

People v. King[10] involved a defendant who was accused of arson in the Watts riot. The initial tape and spectrogram of his voice were obtained from a television news interview in which the defendant admitted the crime; the second spectrograph was prepared from a surreptitiously taped interview with the police. An intermediate appellate court, after a brief but excellent review of the theory and experimental data, held that spectrographic voice identification evidence was inadmissible,

5. 17 U.S.C.M.A. 183, 37 C.M.R. 447 (1967).
6. *See* discussion of experts in relation to admission of such evidence, in text accompanying notes 57, 58 & 62 *infra*.
7. 17 U.S.C.M.A. at 189; 37 C.M.R. at 453.
8. 49 N.J. 343, 230 A.2d 384 (1967).
9. *See* note 28 *infra* and material cited therein for a brief description of this experiment.
10. 266 Cal. App. 2d 437, 72 Cal. Rptr. 478 (1968).

since the technique's accuracy had not achieved sufficient scientific acceptability.

Since *King*, the instant case is the only reported case to deal with the issue.[11] To evaluate whether the admission of this evidence meets the standard for scientific evidence,[12] it is necessary to explain spectrography, the theory, the machine used and the experimental data on which the court relied.

III. THE TECHNIQUE OF SPEAKER IDENTIFICATION

A. Human Speech

English speech, consisting of approximately 40 vowel and consonant sounds, is produced by the passage of air through the vocal folds and the modification of the air stream by the tongue, teeth and lips. Vowel sounds depend on the height of the tongue and the size of the oral cavity, while consonants are dependent on the manner in which the air stream is modified as it passes through the mouth. Organic differences among individuals combined with the individualized process of learning to speak result in identifiable differences when separate individuals utter the same speech sound. This difference is known as interspeaker variability and is recognized as a common phenomenon.[13]

The existence of interspeaker variability raises the theoretical possibility of identification of speakers. However, this identification is somewhat limited by *intra*speaker variability:[14] a difference in one speech sound uttered by one individual from time to time. Some of the common factors accounting for intraspeaker variability include psychological stress, aging and disease.[15] An additional factor, contributing to both inter- and intraspeaker variability, is the phonetic context, isolated or con-

11. However, apparently the use of spectrographic voice identification evidence is growing. One of the few experts on spectrographic voice identification reports that such evidence has recently been admitted in six cases. Paper by Dr. Oscar Tosi, Mich. State Univ., presented before the Acoustical Society of America in Jan., 1972 [hereinafter cited as Tosi Paper].

12. 2 J. Wigmore, Evidence, §§ 222, 411-14 (3rd ed. 1940); 3 J. Wigmore, Evidence, §§ 795, 795a (3rd ed. Chadbourne rev. 1970). For discussion of the legal standard for admission of scientific evidence, see text accompanying notes 52 *et seq. infra*.

13. M. Hecker, *supra* note 2, at 4-5, 13.

14. *Id.* at 16-18. This difference is not completely understood. MSU Experiment, *supra* note 2.

15. M. Hecker, *supra* note 2, at 16-18; Endres, Bambach & Flosser, *Voice Spectrograms as a Function of Age, Voice Disguise, and Voice Imitation*, 49 J. Acoustical Soc. Am. 1842 (1971).

nected, in which the sounds are uttered. An isolated context occurs when one sound or word is spoken alone, whereas the connected context occurs in continuous speech where each sound or word is affected by what came before and after it.[16] While intraspeaker variability and the context of speech complicate the process, interspeaker variability is sufficiently great to make identification possible.[17] There are two main methods of voice identification: listening and visual inspection of a spectrogram. The first is legally acceptable[18] and the second was accepted in *Trimble* for limited purposes.

B. THE SPECTROGRAM PROCESS

> The sound spectrograph consists of four basic parts: (1) a magnetic recording device, (2) a variable electronic filter, (3) a drum which is coupled to the magnetic recording device and carries a sheet of special paper, [sensitive to] . . . (4) an electric stylus which marks the paper as the drum rotates. The magnetic recording device is first 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 [*e.g.* 2.4 seconds]. The speech sample is then played back over and over again 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 this 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 pass-band of the variable electronic filter moves to increasingly 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.[19]

The pattern's dimensions are: (1) time—the horizonal axis, (2) frequency, *i.e.*, pitch—the vertical axis, and (3) amplitude, *i.e.*, loudness—the darkness of the vertical, frequency lines.[20]

The process of spectrogramic identification requires the use of two spectrograms: first, there must be an original, which

16. M. HECKER, *supra* note 2, at 12-13.
17. *Id.* at 18.
18. 2 J. WIGMORE, EVIDENCE §§ 222, 660 (3rd ed. 1940). The available research indicates that listening is a more accurate system of identification than spectrographic voice identification. M. HECKER, *supra* note 2, at 71-73; Stevens, Williams, Carbonell & Woods, *Speaker Authentication and Identification: A Comparison of Spectrographic and Auditory Presentations of Speech Material*, 48 J. ACOUSTICAL SOC. AM. 1596 (1968) [hereinafter cited as Stevens].
19. M. HECKER, *supra* note 2, at 50-51.
20. *Id.*

in the instant case would have been produced from the luring telephone call; and second, an investigatory spectrogram, which in the instant case was produced from the telephone conversation arranged through the Welfare Department. Of the various tests used in spectrography,[21] the one best suited to criminal investigation is the discrimination test in which the examiner's task is to decide whether the original and the investigatory spectrograms represent the same speaker. The test is carried out in the following manner.

> [T]he two recordings are carefully transcribed to facilitate the search for suitable cue material. In the selection of cue material, if there are only a few words common to both recordings, it may be necessary to include phonetically identical portions of different words. Pairs of spectrograms are then prepared for all selected cue materials. The observer examines each pair of spectrograms and determines the degree of similarity of the spectral features. [The examiner may decide] either the two recordings are ascribed to the same speaker, they are ascribed to different speakers, or the results are considered inconclusive.[22]

Several variables are known to affect the ability to identify a speaker by means of spectrograms and some of these are the same as the sources of inter- and intraspeaker variability. The following discussion is limited to those variables which seem significant in the criminal context. The first variable is the selection of cue material, which consists of the specific words used.[23] The second variable is the context of the cue material, isolated or connected.[24] The quality of the transmission of the speech signal may also affect the ability to identify. This variation in quality may be caused by the presence of outside noise in the system or from distortion of the speech signal.[25] A final source of variability is the qualification of the examiner.[26]

C. The Recent Experimental Evidence

Prior to *King*,[27] the scientific community dealing with spec-

---

21. The other tests are the multiple-choice identification and the identification-discrimination tests, which are discussed in M. Hecker, *supra* note 2, at 66, 68.

22. *Id.* at 66-67.

23. *Id.* at 58-60.

24. Isolated and connected context are defined in the text following note 15 *supra*, and in note 34 *infra*.

25. M. Hecker, *supra* note 2, at 61-62. Distortion of the speech signal may arise from the limitations of the electronic equipment used to obtain the samples, *e.g.*, tape recorder and telephone. *Id.* For instance, the telephone passes only certain frequencies.

26. *Id.* at 63-65.

27. People v. King, 266 Cal. App. 2d 437, 72 Cal. Rptr. 478 (1968).

trographs generally felt that there was insufficient information to warrant accepting the spectrographic identification technique as accurate enough for legal proceedings.[28] There was a call for more experiments designed to supply information.[29] At that time the available information indicated that identification by listening was more accurate than spectrographic identification.[30]

Since then an experiment has been conducted at Michigan State University,[31] the results of which were cited with approval in the instant case.[32] Several important variables known to affect inter- and intraspeaker variability and the process of spectrographic identification were accounted for in the experiment. They included the number of cue words used,[33] the number of times each cue word was uttered by each speaker, the quality of transmission of sound, which varied from quiet to noisy, the context in which the cues were uttered,[34] the time elapsed between the recording of the original and investigatory cues and the use of open and closed trials.[35]

---

28. M. HECKER, *supra* note 2; Bolt, Cooper, David, Denes, Pickett & Stevens, *Speaker Identification by Speech Spectrograms: A Scientists' View of Its Reliability for Legal Purposes*, 47 J. ACOUSTICAL SOC. AM 597 (1970) [hereinafter cited as Bolt]. While this article contains a specific statement that it does not represent the views of the Acoustical Society of America, it is generally accepted by members of the profession as representing the Society's view. Interview with Charles E. Speaks, Professor of Communication Disorders, University of Minnesota, in Minneapolis, Jan. 25, 1972. In *Trimble* the court accepts this study as the position of the Society. 192 N.W.2d at 435.

The view that spectrographic voice identification is accurate was initially advanced by Mr. Kersta, on the basis of his studies described in Kersta, *Speaker Recognition and Identification by Voiceprints*, 40 CONN. B.J. 586 (1966), and in State v. Cary, 99 N.J. Super. 323, 239 A.2d 680 (Sup. Ct. 1968), *aff'd per curiam*, 56 N.J. 16, 264 A.2d 209 (1970). These experiments involved closed trials with fixed context statements, and in this connection see text following note 37 *infra*. These studies demonstrated an accuracy rate in excess of 99 per cent; however, their utility in the criminal context is doubtful. M. HECKER, *supra* note 2, at 69.

29. Bolt, *supra* note 28.

30. M. HECKER, *supra* note 2, at 72; Stevens, *supra* note 18.

31. MSU EXPERIMENT, *supra* note 2.

32. 192 N.W.2d at 438.

33. MSU EXPERIMENT, *supra* note 2, at 6. The words used were: it, is, on, you, and, the, I, to, me.

34. *Id.* at. 6. Three contexts were used: isolated, in which the cue words were spoken alone; fixed, in which the cue words for both the original and investigatory spectrograms were uttered in connected speech in the same sentence; and random, in which the same cue words were uttered in connected speech in different sentences for the original and investigatory spectrograms.

35. *Id.* at 7. A closed trial is one in which the examiner knows that one of the original spectrograms matches the investigatory spectro-

The speakers used in the experiment were 250 college age American males, without speech defects, selected at random from a population of 25,000 Michigan State University students. Two of the experimental conditions are important. The spectrographic examiners did not listen to the taped cue material; they visually observed the spectrograms. Second, the examiners were required to make an identification decision, match or no match, in each trial; however, the examiners subjectively rated the certainty of their decisions.[36]

For purposes of determining whether the MSU Experiment results demonstrate that the spectrographic voice identification technique meets the standard for admission of scientific evidence, only those trials with incidents analogous to the common criminal case are significant.[37] These are open trials involving non-contemporaneous utterances in a fixed or random context.[38] The open trial is one in which the examiner does not know that a match exists between an original and an investigatory spectrogram. The non-contemporaneous feature is required because normally the investigatory cue material will be acquired by the police at some time after the original cue material. And the random or fixed context is necessary because it is unlikely that either the original or investigatory cue material would be spoken in isolation.

Certainly the number of errors in a process is indicative of its accuracy. In the MSU Experiment the examiners could make three types of errors in identification: (1) a misidentification—incorrectly matching an investigatory spectrogram with an original spectrogram; (2) a false identification—matching spectro-

---

gram. An open trial is one in which the examiner is not aware whether or not there is a match between the investigatory and original spectrograms. Two thirds of the trials in the MSU Experiment were open trials.

36. Another condition existing in the MSU Experiment was that there were three groups of examiners, composed of women, college students and criminal justice studies students. Each group was broken into panels of various sizes. No difference in performance was noted among the groups but the size of the panels affected accuracy. This was attributed to group dynamics. MSU EXPERIMENT, *supra* note 2, at 10, 14.

37. However, the other trials are important since one of the purposes of the MSU Experiment was to replicate the results obtained by Kersta (*see* note 28 *infra*), and this result was achieved. MSU EXPERIMENT, *supra* note 2, at 15, 19.

38. In *Trimble*, the court notes and was apparently impressed with the fact that the MSU Experiment consisted of approximately 34,000 trials. 192 N.W.2d at 438. However, the number of trials with the incidents common to criminal investigations was substantially less.

grams when no match exists; and (3) failure—missing an existing match.[39] In the criminal context, the first and second types of errors would produce indentification of an innocent person while the third type would fail to identify a guilty person. The MSU Experiment was conducted in two cycles. Table 1 shows the percentage of errors made in the trials. The total range of errors is from approximately 14% to approximately 18%, of which approximately one-third were types (1) and (2), while the remaining errors were of type (3).

Table 1[40]

		Percent Error	
Cue Context	Type of Error	Cycle I	Cycle II
	1, 2, 3	18.26	15.10
	1, 2	6.43	4.81
Random	3	11.83	10.29
	1, 2, 3	14.35	14.84
	1, 2	4.22	4.27
Fixed	3	10.13	10.29

The experimenters made several comments on factors they believed would reduce the errors of types (1) and (2). The examiners spent about 15 minutes on each identification trial and other research indicates that as the amount of time allowed the examiner is increased the rate of accurate identification increases.[41] The experimenters, relying on other experimental data, hypothesized that by allowing the examiners to listen to the tapes of the cue material as well as visually inspect the spectrographs the rate of errors could be reduced.[42] An analysis of the certainty ratings by the examiners convinced the experimenters that allowing the examiners the choice of no decision would reduce the error rate.[43] The experimenters concluded that with these adjustments the error rate could be reduced

39. MSU Experiment, *supra* note 2, at 11. It should be borne in mind that in each experimental trial in the MSU Experiment, the examiner has 10, 20 or 40 original spectrograms; this would not correspond with the usual criminal investigation.

40. *Id.* at table 6.

41. *Id.* at 20; M. HECKER, *supra* note 2, at 72; Stevens, *supra* note 18.

42. MSU EXPERIMENT, *supra* note 2, at 20.

43. *Id.* This would be the normal procedure in a spectrographic identification, described in the text accompanying note 22 *supra*.

to about two per cent.[44] Subsequently, the principal experimenter proposed that admission of spectrographic voice identification evidence be conditioned upon the existence of these accuracy enhancing factors.[45] Even a long-time chief critic of the admission of spectrographic evidence modified his position on the basis of the experiment described.[46] This critic now concludes, with certain limitations,[47] that the per cent of type (1) and (2) errors ranges from 3% to 6%,[48] and concludes with the accuracy of the system established, admission of the evidence is to be determined by the legal standard.[49]

IV. THE LEGAL STANDARD

When scentific evidence is presented at trial two distinct issues arise: the qualification of the process and the qualification of the experts who have used the process.[50] The qualification of the process is a two step procedure: first, the general process must be qualified as accurate; and second, the specific use of the process in the instant case must be established.[51] The standard for qualifying the general process is its general acceptance within the relevant scientific community.[52] The standard which may be satisfied by expert testimony, scientific au-

44. MSU EXPERIMENT, *supra* note 2, at 20.

45. Tosi Paper, *supra* note 11, at 2-3.

46. Letter from Dr. Peter Ladefoged, Professor of Phonetics, U.C.L.A., Feb. 17, 1972 [hereinafter cited as Ladefoged Letter]. Dr. Ladefoged testified for the defense in Cary, King and the instant case.

47. Dr. Ladefoged limits his acceptance of the MSU Experiment results to cases not involving women, as in the instant case, mimics or disguised voices. Ladefoged Letter, *supra* note 46, at 1-2. *Compare* Endres, Bambach & Flosser, *Voice Spectrograms as a Function of Age, Voice Disguise, and Voice Imitation*, 49 J. ACOUSTICAL SOC. AM. 1842 (1971).

48. Ladefoged Letter, *supra* note 46, at 1-3. The reason for the range is that Dr. Ladefoged, while agreeing that the accuracy enhancing factors noted in the text would reduce the MSU Experiment errors, suggests that the presence of "confusable voices" might increase errors. A confusable voice is one which sounds the same as another. *See* M. HECKER, *supra* note 2, at 57. For an example of what may be a confusable voice and an example of an erroneous spectrographic identification, see New York Times, March 27, 1971, at 57, col. 2; Wall Street Journal, March 13, 1972, at 1, col. 1.

49. Ladefoged Letter, *supra* note 46, at 3.

50. 2 J. WIGMORE, EVIDENCE § 414 (3rd ed. 1940).

51. J. WIGMORE, THE SCIENCE OF JUDICIAL PROOF § 220 (3rd ed. 1937).

52. Frye v. United States, 293 F. 1013 (D.C. Cir. 1923); C. MCCORMICK, EVIDENCE § 170 (1954). This rule has been adopted in Minnesota, at least in relation to the lie detector. State v. Perry, 274 Minn. 1, 142 N.W.2d 573 (1966); State v. Kolander, 236 Minn. 209, 52 N.W.2d 458 (1952).

thors or judicial notice,[53] would not seem to be reasonably open to challenge[54] since it appears to be the only workable system which allows advances in knowledge to assist in accurate adjudication, and yet at the same time protects the parties from the effects of unfounded speculation. Whether there is acceptance of the process by the relevant scientific community is a matter of law.[55]

The acceptance standard has not established how large the relevant scientific community must be. In the field of spectrographic voice identification, there are no more than a handful of individuals who have testified in the reported cases.[56] However, this paucity of present experts alone should not bar admission. The acceptance standard should be satisfied if the existing experimental data has been evaluated and accepted by a sizable number of the relevant scientific community though each member himself may not be a qualified expert in the particular scientific process. To date, it appears that only the two experts who testified in the instant case have reached a consensus on the level of accuracy of the spectrographic voice identification technique.[57] This small acceptance would seem to be insufficient to show acceptance by the scientific community since there are surely others, in speech, speech therapy, audiology and related fields, who would be qualified to assess the results of the MSU Experiment. More scientific evaluation, comment and replication of the MSU Experiment are needed as a means of further verifying the accuracy of the spectrographic voice identification technique.

When this scientific acceptance has been demonstrated then it will be for the court to determine whether the accepted level of accuracy is sufficient to warrant admitting evidence based on

53. 2 J. WIGMORE, EVIDENCE § 414 (3rd ed. 1940).

54. It has been suggested with particular reference to spectrographic voice identification, that accuracy alone should be the test. Note, *Evolving Methods of Scientific Proof*, 13 N.Y.L.F. 679, 745-51 (1968); Note, *Evidence—Voiceprint Methods of Identification—Reluctance of the Courts Toward Acceptance of Scientific Evidence*, 12 N.Y.L.F. 501 (1966). But the problem, as Wigmore resolves it, is one of how to determine accuracy. *See* note 53 *supra*.

55. C. MCCORMICK, EVIDENCE § 53 (1954); 2 J. WIGMORE, EVIDENCE § 561 (3rd ed. 1940). *But see* C. MCCORMICK, EVIDENCE § 170 (1954).

56. These include L. Gerstman, City College, C.U.N.Y., F. Clarke, Stanford Research Institute, V. Fromkin, U.C.L.A., P. Ladefoged, U.C.L.A., O. Tosi, Michigan State University, and E. Nash, Michigan State Police.

57. These were Dr. O. Tosi and Dr. P. Ladefoged.

the scientific process. As noted above,[58] the type (1) and (2) error rate in the MSU Experiment is in the range of six per cent. The experimenters claim that this would be reduced to approximately two per cent by requiring that the examiners listen to the tapes from which both spectrograms are prepared as well as visually inspect the spectrograms, that they be given as much time as they desire and that they be given the choice of no decision. It must be emphasized that this two per cent figure is only an estimated ideal; it has *not* been achieved. Additional research, accounting for these supposed accuracy-enhancing factors should be conducted and it should attempt to discover whether, with these factors, spectrographic voice identification is more or less accurate than aural identification alone. If such research could demonstrate that an error rate of only two per cent could be normally expected, then a court would seem justified in approving the general process of spectrographic voice identification.

V. CRITICISM OF TRIMBLE

The *Trimble* court never squarely faces the issue of generally qualifying the process of spectrographic voice identification. With respect to the first step of that qualification,[59] the court ambiguously quotes one expert as saying that Dr. Tosi's results (in the MSU experiment) have been "accepted by the scientific community with certain limitations."[60] One of these limitations was that the process was not as accurate for females, but this fact apparently did not pose any difficulty for the court. The second step, determining the accuracy of the process, is briefly mentioned without any analysis.[61] Instead, the court subsumes the question of qualifying the process of spectrographic voice identification under the question of qualifying the expert. The court reasons:

---

58. *See* text following note 39 *supra.*

59. *See* text accompanying note 51 *supra.*

60. 192 N.W.2d at 440.

61. The only time the court refers to the accuracy of the process of spectrographic voice identification is when it quotes from Dr. Tosi's testimony in which he responded to the question of how reliable is an examiner's opinion given certain ideal conditions. Dr. Tosi answered,

> Providing that all these conditions that you have expressed, especially that the examiner is responsible and he is allowed to say, "Well, I don't know, I cannot produce in this case an identification," and only in those cases in which he is absolutely sure of his statement, I think that then the method is very highly extremely reliable.

Id. at 439.

> It is common knowledge that the opinion of an expert on an identification subject is seldom so infallible that others in his field do not disagree with him. But disagreement alone does not make the opinion inadmissible. Where experts disagree, it is for the factfinder, whether that be jury or court, to determine which is more credible and therefore more acceptable.[62]

In short, the court takes an almost passive role regarding admissibility, leaving it to the jury to decide what "weight" to give the evidence.

The court specifically concludes that spectrograms should be admissible at least for the purpose of corroborating identifications by aural voice comparison. It reasons that since spectrograms have been shown to aid aural identifications, which themselves are admissible, that spectrograms should "be admissible at least for the purpose of corroborating opinions as to identification by means of ear alone," and for impeachment.[63]

The court's conclusion is unacceptable for a number of reasons. First, despite the recent MSU experiment,[64] the process of spectrographic voice identification does not appear to possess either the requisite scientific community acceptance or sufficient accuracy to warrant admissibility.[65] Second, limiting doubtfully accurate evidence to corroborative purposes will most likely not insure that the infirmities inherent in such evidence will not in fact be realized. It is likely that a jury will not confine scientific evidence to such a limited role, but will also use it as substantive evidence. Further, there appears to be no precedent in the area of scientific evidence for such a limitation.[66] If the court was not convinced of the accuracy of spectrographic voice identification, then the proper resolution would have been to deny admission, rather than limiting its use and relegating the debate over its acceptance by the scientific community to a question of weight for the jury.

---

62. *Id.* at 440.
63. *Id.* at 441.
64. MSU EXPERIMENT, *supra* note 2.
65. *See* Part IV *supra.*
66. Generally, corroboration is a requirement for and not a limitation on the admissibility of evidence. Classic instances requiring corroboration include treason, testimony of accomplices and rape. *See generally* 7 J. WIGMORE, EVIDENCE §§ 2032-75 (3rd ed. 1940).