5-1-1978

Voiceprints

D. M. Stotland
G. O. Brown

Follow this and additional works at: https://digitalcommons.schulichlaw.dal.ca/dlj

Part of the Courts Commons

Recommended Citation
I. Introduction and General Background

The use of the voiceprint technique of speaker identification was first suggested in 1962 by a Mr. Lawrence Kersta who was at that time a worker at the Bell Research Laboratories in the United States of America. Since that time, the technique has been subject to a great deal of legal and scientific controversy in the United States. The desire of the authors to write this article stems from the fact that very recently the technique has been considered twice by courts in Canada, albeit only at the trial level. In both cases expert opinion based on voiceprint analysis was held admissible at least insofar as it was corroborative of identification by ear. The authors’ purpose in this article is to examine the technique, the scientific studies and criticisms of it, the U.S. appellate decision and Canadian case law to date; finally, we would like to put forward what we feel should be the course taken by a Canadian appellate court.

It is perhaps obvious that speaker identification is often a crucial point in a modern criminal trial. If not, then consider the words of Detective Lieutenant Ernest W. Nash of the Voice Identification Unit of the Michigan State Police (and recently elected senator for Michigan), a recognized expert in the area of voiceprint analysis for speaker identification:

As an instrument of crime, the human voice is used more frequently than an auto and weapon combined. . . . There are certain crimes due to the nature of which the offender’s voice is the only evidence to prove

*D. M. Stotland, LL. B. Dalhousie, 1977;
**G. O. Brown, LL. B. Dalhousie, 1977

The authors wish to thank the following people for their invaluable assistance in this study: Professor G. Duncan for numerous useful references and valuable suggestions; Professor C. Hutchins for his comments and references in Canadian evidence law; Corporal K. Taylor for portions of the transcript of R. v. Medvedew which was generously supplied gratis; and Mr. L. R. McInnis, Crown Attorney of Manitoba for his valuable oral and written comments.

2. Regina v. Montani (1974), 26 C.R.N.S. 339 (Ont. Prov. Ct.) and Queen v. Medvedew (1976), Queens Bench Assize Court, Brandon, Manitoba, as yet reported (this case is on appeal)
the crime and connect the offender with it. There are others in which the voice is the most valuable evidence to identify the offender with his crime.\footnote{3}

It is then fairly obvious that if possible it would be highly desirable to have a scientifically accurate and reliable means of speaker identification. In 1962 Lawrence Kersta put forward a claim that he had found such a means. Prior to 1966 all voice identifications were made by the human ear, by someone familiar with the sound of the voice being identified. Although generally accepted by the courts, it has been recognized that such identifications are occasionally quite unreliable.\footnote{4}

However, in 1966 in the case of \textit{People v. Straekle}\footnote{5} the New York Trial Court admitted evidence of voiceprint analysis made by Kersta for the purposes of identification. This was the first case anywhere to adjudicate on the admissibility of spectrographic analysis of voice recordings as a means of identification and it marked the commencement of a controversy which remains unresolved to this day. However, before going on to a more detailed examination of the technique it would be appropriate to give a brief and simple explanation of just what a voiceprint is.

A voiceprint, more properly known as a voice spectrogram, is the output of a machine known as a sound spectrograph. This machine will take a taped segment of speech and transform the acoustical information on the tape into a visual display — the voiceprint.\footnote{6} Of the two basic forms of spectrograms producible from a given voice sample, bar and contour, the format most frequently used in voice identification is the bar spectrogram.\footnote{7} A copy of a typical bar spectrogram is shown below.\footnote{8}

\begin{quote}
\textbf{References}
\begin{itemize}
\item[3.] Ernest W. Nash, \textit{Voice Identification by the Voiceprint Method} (Michigan State Police Handbook), at 1. Nash cites the following crime categories as examples: murder, attempted rape, rape, kidnapping, extortion, abortion, bribery, burglary, unarmed robbery, sale of narcotics, conspiracy to gamble, larceny, unauthorized use of aircraft, impersonating a police officer, bomb threats, false fire alarms, false report of crime, threatening phone calls, obscene phone calls, and nuisance phone calls.
\item[4.] Bernard S. Kamine, \textit{The Voiceprint Technique: Its Structure and Reliability} (1969), 6 San Diego L. Rev. 213 at 213
\item[5.] Crim. No. 9326/64 (Sup. Ct. Westchester Cty., N.Y. 1966)
\item[6.] David M. Seigal, \textit{Cross-Examination of a "Voiceprint" Expert: A Blueprint for Trial Lawyers} (1976), 12 Crim. L. Bul. 509 at 510
\item[7.] William R. Jones, \textit{Evidence vel non The Non Sense of Voiceprint Identification} (1974), 62 Kentucky L. J. 301 at 303
\item[8.] \textit{Supra}, note 3 at 3
\end{itemize}
\end{quote}
On a bar spectrogram, the time dimension is plotted from left to right, i.e., the beginning of the voice segment being analyzed is at the left, and the end is at the right. Frequency is plotted along the vertical axis with the lower frequencies at the bottom. Intensity is exemplified by the density of the lines on the spectrogram—the darker the lines, the greater the intensity of the sound at that particular frequency and time.9

II. The Voiceprint Technique of Speaker Identification

1. The Sound Spectograph

What the ear perceives as sound is basically nothing more than the back and forth vibration of air molecules, commonly called sound waves.10 Even the most complex of sounds is in reality just a superposition of a number of simple component waves each having its own frequency and amplitude. Speech consists of such complex, multi-frequency sounds.11 The sound spectrograph is quite simply a machine designed to break down complex sounds into their component frequencies.12

9. Supra, note 7 at 303
10. Supra, note 4 at 218
11. Id. at 219
12. Id. at 218
The primary component of the machine is the analyzing filter.\textsuperscript{13} As this is not a paper in electrical engineering, we are restricting our discussion of this filter to a "black box" treatment. In effect the filter operates as follows: for a given complex, multi-frequency signal put into the filter, the filter will allow only those component frequencies of that input signal which come within a narrow band of frequencies to pass through it. The width of this "pass" band of the filter is known as the filter displacement.\textsuperscript{14} The filter is variable in that the position of the "pass" band on the frequency scale may be changed by making adjustments to the filter (analogous to tuning a radio).

The other major components of the spectrograph are the stylus and the electrically sensitive recording paper.\textsuperscript{15} The analyzing filter is coupled to the stylus controlling the electric current which is sent through the stylus. The current varies in power as does the amplitude of each particular component frequency of that complex sound. For each component frequency this current makes a dot on the facsimile paper which varies in darkness according to the power of the current, thereby indicating the relative amplitude of the particular component frequency.

The typical paper has a 2:1 ratio, \textit{i.e.}, if a certain amplitude causes the faintest possible dot, an amplitude twice that power makes the darkest possible dot.\textsuperscript{16}

In order to get a complete spectrogram the taped segment of sound being analyzed is first wrapped around an analyzing drum on the spectrograph. The drum is then rotated continuously and for every revolution the sound being analyzed is scanned and the complex, multi-frequency signal is fed to the analyzing filter.\textsuperscript{17} Initially the "pass" band of the filter is set at the bottom of the frequency scale, but with every replay of the tape it is moved up the frequency scale by an amount equalling the filter displacement.\textsuperscript{18} The facsimile paper is similarly mounted on a synchronously rotating recording drum and each time the tape is replayed and the filter moves to a new displacement level, likewise the stylus moves.

\textsuperscript{13} \textit{Id.}
\textsuperscript{14} \textit{Id. at 220}
\textsuperscript{15} \textit{Id. 219}
\textsuperscript{16} \textit{Id. 220}
\textsuperscript{17} \textit{Supra, note 7 at 303}
\textsuperscript{18} \textit{Supra, note 4 at 220}
correspondingly up the frequency scale. In the end the facsimile paper is covered with a series of scan lines and we have a complete spectogram of the taped segment of sound.

Some aspects of the spectrograph will be brought out in greater detail later in this section.

2. Speech Production

Some knowledge of the acoustics of speech production is essential to a proper understanding of the voiceprint technique of speaker identification. By the very nature of the human organs of speech, the acoustics of speech production must be described in terms of resonating air-filled chambers or cavities and the associated effects of reinforcement, production of subsidiary frequencies, and damping.

Any body which is capable of being elastically deformed can be forced to vibrate. However, any such body has what is known as a natural frequency. This natural frequency is determined solely by the physical constants of the body. The interesting thing about the natural frequency of such a body is that this is the frequency at which it seems to “like” to vibrate (i.e., it takes much less energy input to get an object to vibrate at its natural frequency than at any other frequency). A simple example of this is the rocking chair. Try rocking the chair at twice the frequency it “wants” to rock.

Now consider a volume of air in a container. This will have a natural frequency depending only on its physical constants. When struck by sound waves of its natural frequency it will tend to vibrate sympathetically at the same frequency (i.e., resonate). However, in such a case not only will the air merely resonate; it will increase the amplitude of the original sound and make it louder (i.e., reinforce it). Anyone who has ever heard the effect of placing the struck stem of a tuning fork on a hollow box can attest to this. Also, in such a case

the air chamber will resonate not only the single tone which activates its natural frequency, but due to its imperfections as a resonator, it will also produce subsidiary sound waves clustered around that natural frequency. These subsidiary frequencies diminish in power the farther they are from the natural frequency

19. Supra, note 7 at 303
20. Supra, note 4 at 222-224
21. Id. at 222
22. Id.
This effect is called damping. It is highly relevant to the voiceprint because the pattern made by damping is governed largely by the shape of the particular resonator. For example, even though two chambers vibrate at exactly the same natural frequency, the degree of the decline in power of the subsidiary sound frequencies will differ if the shapes of the two chambers differ.\textsuperscript{23}

Before going on to consider exactly how speech is produced and how this relates to what we see on a spectrogram there is one more term which must be explained — tones. If all the frequencies in a certain sound are multiples of a single frequency they are called tones (note: if the frequencies are not related in this manner they are called noise).\textsuperscript{24}

Ignoring whispers and certain consonants, the glottal tone produced by the passage of air through the vocal chords is the originating point of all speech.\textsuperscript{25}

As this tone passes through the vocal tract, it is resonated, reinforced, and damped in the cavities of the nose, and throat, which are of fixed shape, and the mouth. Although the cavities formed in the mouth are pre-determined to some extent by the anatomy of the individual, they are largely created by the movement of the tongue, the jaw, and the lips — the articulator.\textsuperscript{26}

Variations in the placement of the articulators are responsible for most of the differences between the various speaking sounds. The process of learning to talk is essentially that of learning how to position one’s articulators so as to achieve a desired sound.\textsuperscript{27}

The size and shape of the various cavities determine the glottal tone frequencies which are resonated, reinforced and damped and the degree of each process. The resulting clusters of relatively high amplitude frequencies which are produced by each cavity are called formants and are reflected on the spectrogram by dark bands called resonance bars . . . . The blackened bar along the bottom of the spectrogram represents the fundamental frequency of the glottal tone, those above represent the formants selected and produced by the various cavities in the head. It is important to note two things: (1) since the sounds from one cavity affect those from the others, resonance bars cannot be attributed solely

\textsuperscript{23} Id.
\textsuperscript{24} Id.
\textsuperscript{25} Id.
\textsuperscript{26} Id.
\textsuperscript{27} Id.
to individual cavities; (2) since the cavities will only resonate their natural frequency and its subsidiaries, changes in the pitch of the voice will not affect the position of the resonance bars in the spectrograms, except in the unlikely case that such a change would eliminate completely the particular natural frequency of one of the cavities from the glottal tone.28

3. The Theory of Invariant Speech and the Spectrograph

Two of the assumptions underlying the voiceprint technique of speaker identification are (1) every person's voice is unique,29 and (2) this uniqueness is adequately portrayed in a spectrogram.30 This first assumption is often referred to as the theory of invariant speech.31 An alternative statement of the theory is: "interspeaker variability is always greater than intraspeaker variability."32

The theory's basic premise is that the characteristics of each individual's speech are uniquely determined by his vocal cavities and articulators. Vocal cavity contours and dimensions vary between persons much like any other part of the anatomy; the likelihood of two people having identical cavities seems remote. As for articulators, not only are they subject to anatomical differences, but their use in forming speech is a complex learned muscle manipulation unlikely to be identical in any two people.33

The combination of these factors would seem to render it plausible that every person's speech is relatively invariant and unique.34 However, it should be noted that though positioning of the articulators is habitual:

these habits can be changed and often are to correct speech defects. Whether an individual can deliberately alter one of these habits, either permanently or temporarily, to effect a disguise, is not certain. Whether the habits are so strong that they will not involuntarily yield in some circumstances . . . is not certain. Whether the habits of one individual might compensate for the anatomical differences between himself and another as they both strive to produce the same speech sound is uncertain; however given the number of variables involved it is unlikely. To what

28. Id.
29. Anon., The Voiceprint Dilemma: Should Voices be Seen and Not Heard? (1975), 35 Maryland L. R. 267 at 271
30. Supra, note 4 at 224
31. Supra, note 29 at 271
32. Supra, note 6 at 510
33. Supra, note 29 at 271
34. Supra, note 4 at 226
extent factors such as laryngitis, colds and dentures alter the spectrogram is uncertain.\textsuperscript{35}

The second assumption involved in the voiceprint technique of speaker identification (\textit{i.e.,} that accepting that everyone's voice is unique, the spectrogram adequately displays that uniqueness) is perhaps also open to some doubt. As one author said:

No one questions that the spectrograph depicts accurately and consistently those parameters of speech that it has been designed to graph. But the theory goes further and assumes that those parameters — frequency and intensity plotted across time — are the ones which truly represent the unique elements of individual speech.\textsuperscript{36}

However, another author has pointed out that:

it is probable that the spectrograph visually portrays many of those features which allow a listener to aurally distinguish one voice from another. Moreover, the ear distorts certain sounds; the spectrograph does not.\textsuperscript{37}

4. \textit{Pattern Matching}

There is a third major assumption underlying the voiceprint technique of speaker identification as it exists today. Even if one accepts that every voice is unique and that the spectrogram adequately portrays that uniqueness, the validity of the technique rests on the assumption that an examiner can accurately discern whether two spectrograms "match", \textit{i.e.,} come from the same person.\textsuperscript{38} The technique requires an examiner to compare a spectrogram made from the recorded speech of an unknown speaker (the perpetrator of the crime) with the spectrograms made from the recorded speech of one or more known speakers (the perpetrator may or may not be in the group of known speakers).\textsuperscript{39}

In coming to a decision as to whether any of the known spectrograms "match" the unknown spectrogram the examiner must use a process of pattern-matching.

Pattern-matching is a standard method of establishing identity between two objects. . . . Whenever used, the features of one sample are compared with those of another, and, if a sufficient

\textsuperscript{35} \textit{Id.}  
\textsuperscript{36} \textit{Supra}, note 29 at 271  
\textsuperscript{37} \textit{Supra}, note 4 at 224  
\textsuperscript{38} \textit{Supra}, note 29 at 271  
\textsuperscript{39} \textit{Supra}, note 6 at 510
number of identical features appear on both, identity is confirmed. However, a prerequisite is that each of those features or the particular combination of those features must not appear on non-identical samples.\textsuperscript{40}

Although it is preferable that the samples from the known speakers contain short phrases identical to those in the sample of the unknown speaker,\textsuperscript{41} it is not absolutely essential. In the absence of such identical phrases, comparison can still be made between excerpts of the steady-state portions of speech representing the static position of the articulators in making the same sound.\textsuperscript{42} Having made up spectrograms from the various samples, the examiner then proceeds to:

compare the relationship among the various resonance bars . . .

at the beginning of the time axis. He continues this matching process along the time axis at significant points comparing the frequency/intensity patterns, widths, shapes, slopes, mean frequencies and separations of the bars as well as the frequency/intensity patterns of incomplete resonance bars, the large dark patterns created by noise consonants, and the pattern of the vertical lines, called striations.\textsuperscript{43}

Striations which are relatively far apart generally indicate a low-pitched voice; conversely striations relatively close together generally indicate the voice is high-pitched.\textsuperscript{44} The examiner should probably not consider the duration or rate of speech as this factor may vary with the speaker's setting and emotional state.\textsuperscript{45}

During the matching-process the examiner is looking for points of similarity and ignores differences among the samples.\textsuperscript{46}

If he finds sufficient points of similarity (variously pegged at 16 or 20), he may conclude that the two voice samples were made by the same person. If the operator cannot find an adequate number of points of similarity, he is unable to make a match and cannot say whether the two voices are the same. In some cases, the spectrograms may be so dissimilar that he will be able to state that the two recordings were \textit{not} made by the same person.\textsuperscript{47} There is a

\begin{itemize}
\item[40.] \textit{Supra}, note 4 at 215
\item[41.] \textit{Id.} at 217
\item[42.] \textit{Id.} at 215
\item[43.] \textit{Id.} at 216
\item[44.] \textit{Supra}, note 6 at 511
\item[45.] \textit{Id.}
\item[46.] \textit{Supra}, note 4 at 216
\item[47.] \textit{Supra}, note 7 at 302
\end{itemize}
problem with just arbitrarily fixing the number of points of similarity necessary for positive identification; the number chosen is not related to the size of the speech samples compared, nor to the type of samples compared (i.e., identical phrases or excerpted steady state sounds).\textsuperscript{48}

Moreover, this points of similarity approach is not so objective as might first appear and "examiners readily admit that they rely on a 'general eyeballing' of the spectrograms and other more subjective factors to reach a decision."\textsuperscript{49}

That it is not always an easy matter for an examiner to accurately decide whether two spectrograms were produced by the same person or not can be demonstrated by reference to figure 2.\textsuperscript{50} Shown are four spectrographs of the words 'on you'. A-1 and A-2 are spectrograms of those words as spoken by the person on different occasions, while B-1 and B-2 are spectrograms of those words as spoken by two different people.

\textsuperscript{48} Supra, note 4 at 216-217
\textsuperscript{49} Supra, note 6 at 511
\textsuperscript{50} Supra, note 7 at 305
At this point, some further reference to the spectrogram should be made. Different spectrograms will be produced from the same sample if the machine's filter displacement is changed. Modulating the filter is necessary in order to get a spectrogram which is not entirely black and this affects the relative darkness of the resonance bars and determines their width. Therefore, it is essential that spectrograms to be compared be made on machines with the same level of filter modulation. Most spectrographs have adjustable frequency scales and the adjustment of the frequency scale will affect the appearance of the spectrogram produced. For these reasons and since the technique involves a process of pattern-matching, it is highly desirable that spectrograms being compared should be made on the same machine with careful attention paid to any adjustments made to the machine.

It should also be noted that spectrograms are affected by the circumstances in which the samples of speech are recorded. For example,

51. *Supra*, note 4 at 221
52. *Id.*
53. *Id.* at 221-222
although the human voice produces meaningful sounds which vary from about 80 cycles to about 8000, most of the equipment used to obtain voice samples accepts considerably less than full range, e.g., the telephone only accepts frequencies up to about 3500 cycles. This is analogous to finding only part of a fingerprint.\textsuperscript{54}

III. \textit{Scientific Studies and Criticisms of the Technique}

1. \textit{Kersta's Claims and Research}

The Bell Telephone laboratories made voice spectrograms a practical reality, in 1948. The theory behind the machine had been known for many years and, as early as 1932, plans for such a machine had been published.\textsuperscript{55} This instrument was soon used, in phonetic research, to show differences and similarities in words. By the early 1960's, one Bell employee was claiming that the technique could accurately distinguish a speaker. Mr. Lawrence Kersta, an electrical engineer and physicist at Bell Laboratories for 30 years, published results of scientific experiments, in 1962, which claimed that speaker identifications by voiceprints were 99\% accurate.\textsuperscript{56}

The design, scope and conclusions of Mr. Kersta's experiments are important in that they help explain the early scientific controversy and voice spectrograms' cool reception in the courts from 1965 to 1969.

Kersta's technicians (high school girls with one week training) compared the spectrogram of an unknown speaker with sets of five, nine or twelve spectrograms of known voices. The examiner had to match the unknown with one of the known in the "closed" set, and thus make an "identification". The sample-set always had a sample of the unknown speaker in it, and voiceprints were made of single isolated words and not words in the context of sentences.\textsuperscript{57} The ten words preferred in the voiceprint field and used in this study were "the", "to", "and", "on", "is", "you", "I", "it", and "a".\textsuperscript{58}

\textsuperscript{54} \textit{Id.} at 222. For a discussion of the technique of recording voice samples which deals with the problems of telephone conversations see, \textit{supra}, note 3 at 10-15. See also text accompanying footnote 92, \textit{infra}.

\textsuperscript{55} R. H. Bolt \textit{et al.}, \textit{Speaker Identification by Speech Spectrograms: A Scientists' View of its Reliability for Legal Purposes} (1970), 47 J. Acoust. Soc. Am. 597 at 604 (hereinunder referred to as "Bolt \#1")

\textsuperscript{56} L. G. Kersta, (Dec. 29, 1963), 196 Nature 1253

\textsuperscript{57} E. W. Clark, \textit{Voiceprints — The admissibility Question: What Evidentiary Standard Should Apply?} (1975), 19 St. Louis U. L. J. 509 at 512

\textsuperscript{58} \textit{Supra}, note 7 at 303
All voiceprints were recorded within a short period of time, so that changes as a function of time were not measured.

Kersta claimed a 99% success rate and stated that the rate of misidentification varied from 0% for the word “a” to 1.8% for the word “to”. He further stated, while giving evidence in People v. King, that the voiceprint technique was as infallible as fingerprints, which was an inapt, rash statement which has plagued more credible voiceprint advocates to this day.

2. Criticisms of Kersta’s Work

Soon after Kersta had published his results he was taking the stand — as an expert — and advocating the use of voice spectrograms as evidence in criminal trials. This action prompted the Acoustical Society of America to set up a “Technical Committee on Speech Communications” to review Kersta’s studies. In 1970, this committee released their findings which strongly criticised Kersta’s work. This “Bolt” study found:

1) That Kersta’s tests merely stated results and conclusions but did not describe the experimental method used. Hence, other research workers could not repeat Kersta’s work, to check his results. The committee also questioned Kersta’s qualifications and expertise.

2) A heterogeneous sample of voices were used for the trials; the use of different accents, ages (voices change with age) and backgrounds was not taken into account. Obviously, if a subject speaks with an accent, his voice spectrograms will be markedly different from the rest of the samples, and identification becomes that much easier.

3) The most vociferous criticism was made over the “closed” design of the tests in which each unknown voiceprint which the examiners were asked to identify had an actual match in a library of voiceprints whose speakers were known. The

59. R. Barnett, Voiceprints: The End of the Yellow Brick Road (1974), 8 U. of San Francisco L. R. 702 at 706
60. (1968), 72 Cal. Rptr. 478 (2d Dist, C.A.)
62. Bolt #1, supra, note 55 at 604 et seq.
63. Supra, note 57 at 512-513
64. W. A. Bacon, Voiceprint Identification: The Trend Towards Admissibility (1974), 9 New England L. R. 419 at 420
examiner's task therefore was merely to identify the "best" match. This procedure failed to account for a feature crucial to the forensic use of the voiceprint technique: the possibility of no match.\(^6\)\(^5\) The examiner may approach the problem by a process of elimination, when he has to in a closed design experiment. In a forensic situation, the situation is like an open design, the examiner can never be sure that the suspect's voiceprint is in his collection of comparative voiceprints. In other words, the examiner \ldots must decide not which is the "best" match but if there is a match at all, since the possibility exists that no one in a group of suspects is the guilty speaker. This element of uncertainty — that there may be no match — renders the forensic or "open" test far more difficult than Kersta's closed tests.\(^6\)\(^6\)

4) Perhaps the most serious criticism was the lack of standardized decision criteria for judging whether one voice spectrograph matched another.\(^6\)\(^7\) The whole process was seen to be highly subjective, and dependant on the skill of the individual examiner. The examiner's decision was the product of a general impression.

Several small studies on voice spectrography followed Kersta's report. The most relevant research paper produced disturbing results. It was reported that the first "open" tests were conducted in 1968.\(^6\)\(^8\) \ldots The tests were "open" in that they explored the possibility of a "no match" result.

In open aural tests only 6-8% of the unknown samples were incorrectly identified with a library match. But in open visual tests, where 87.5% of the unknown speakers actually had a match, there was a 31% incorrect authentication rate, and where only 50% of the unknown speakers actually had a match, there was a 47% rate of incorrect authentication.\(^6\)\(^9\)

It was also found (in that study) that identification of the speaker varied directly with the size of the utterance. The greater the message size the greater the chance of recognition.\(^7\)\(^0\)

65. Supra, note 29 at 272
66. Id.
67. Supra, note 57 at 513
69. Supra, note 29 at 272
70. Id.
In light of the almost universal criticism of Kersta in the literature, it is not surprising that voice spectrograms were accepted as valid scientific evidence in very few courts. The technique might have slowly faded into irrelevance and history had it not been for the 1972 Tosi Report.

3. The Tosi Experiments

Professor Oscar Tosi and five other scientists conducted two years of experiments which were funded by a $300,000 grant from the United States Department of Justice. Professor Tosi was aided by the Michigan State Police through every phase of the study.

Twenty-nine people were selected and trained in voiceprint analysis for one month, (Kersta’s high school girls had had only one week of training) and they eventually completed a total of 34,996 trials of identification in the course of the study.

Each examiner had fifteen minutes to make a forced choice from only visual material. No aural samples were ever examined. The 250 subjects, whose voices were analyzed in various trials, were selected at random from 25,000 American-English speaking male students from Michigan State University. Kersta’s problem of heterogeneity and accents was thus avoided.

Professor Tosi divided the experiment into two major studies. The first study which took up one-third of the trials, consisted of “closed” experiments to test Kersta’s results. Kersta’s results were repeated in this study. The second study examined other types of spectrogram analyses which more closely resembled forensic situations. “Open” experiments were utilized in those trials, and several variables were examined.

1) The number of clue words examined was varied, the first set had six words and the second set was three words greater for a total of nine. The words used were “it”, “is”, “on”, “you”, “the”, “and”, “to”, “I” and “me”. These particular words were selected because of their high frequency in the English language. The difference between the two sets was not found to be significant.

72. Supra, note 57 at 513
74. Id. at 2033
2) The clue word was presented one, two or three separate times from the same speaker. It was found that no significant cognitive improvement was noted upon increasing the amount of exposure.

3) Three differing recording routines were compared, a) direct recording of the voice; b) recording over a telephone in a quiet environment, and c) recording over a telephone in a noisy environment (50 dB/L of white noise). No significant difference in error rate was seen in all three conditions.75

4) The study then examined different contexts of the clue words used for speaker identification. Three types of contexts were tested:
   (i) clue words spoken in isolation;
   (ii) clue words spoken in a fixed context — the same sentences produced by known and unknown speakers were compared; and
   (iii) clue words spoken in a random context — different sentences containing the same clue words were compared.76 There were significant differences found among these three experimental modes and the level of difference varied with several other experimental variables that were tested, such as the number of examined voiceprints and the time interval allowed between voice recordings of the same individual.77

5) The number of known speakers — whose voiceprints were examined to match the unknown exemplar — was varied from 10 to 40. Dr. Bolt analyzed Dr. Tosi's published experimental data and found that the increase “. . . at least doubled the probability of making a false identification”.78 This finding is significant because it is conceivable that a technician may test dozens of voiceprints, on file, in order to identify a suspect's voice recording.

6) The effect of the passage of time between successive voice spectrograms was investigated. In the first, the matching

75. Supra, note 59 at 712-713
76. Supra, note 73 at 2033
77. Id. at 2041
spectrograms were produced at a single, contemporaneous recording session. In the second, the recordings were produced under non-contemporaneous conditions, taken one month apart.  

The passage of time between voice recordings did raise the percentage of error up to about 12%.  

The results of the study were succinctly summarized by Siegel as follows:

The error rate was below one percent where the examiners knew that a match existed (closed test), contemporary matching spectrograms were used, and cue words were spoken in isolation. This result obviously has little, if any, application to the typical forensic setting.

Other error figures reported by Dr. Tosi are more relevant to forensic application of the technique. In open trials using non-contemporary matching spectrograms and clue words spoken in a random context, the total error rate was 18.26 percent, of which 6.43 percent were errors of false identification and 11.83 percent were errors of false elimination.

Professor Tosi explained that the 11.83% error of false elimination would not adversely affect the accused (he would be removed from suspicion). But those errors detract from the technique's reliability, and law enforcement efficiency.

The 6.43% error of false identification was a more serious matter.

Tosi's Rationale

Professor Tosi suggested that certain factors could reduce this error rate to a great degree. He argued that circumstances encountered in actual forensic voice spectrography were very different from experimental conditions. Doctor Tosi asserted that:

1) Expert technicians, with wide experience, would be used instead of student's with only one month of training.

2) In real-life situations where the examiner has more time to make his decision and

3) is not required to make the identification where he is unsure, these error rates would fall to approximately two percent and five percent respectively. Dr. Tosi derived these rates by

79. Supra, note 59 at 712
80. Supra, note 6 at 520
81. Id.
multiplying the actual error rates experienced in the tests (six percent and thirteen percent) by the proportion of wrong comparisons the examiners labeled as “certain” (forty percent)\(^{82}\)

It was found that there was a lower rate of error in decisions when the examiner expressed great confidence about his decision.

4) Trained voiceprint experts do both *aural* and visual examinations which has been shown to be more accurate.\(^{83}\) (the experimental examiners only made visual determinations)

5) A professional voiceprint analyst would be more motivated in his work and thus do a better job, because the accused’s liberty would be at stake and the expert’s job would be lost for poor performance.\(^{84}\)

As convincing as the above reasoning may be it must remain in the realm of conjecture, until it is tested in controlled experiments.

After the Tosi study was published numerous appellate courts ruled that voiceprints were good evidence. There was a hiatus in the criticism from voiceprint opponents as they digested Dr. Tosi’s study. (This silence was interpreted by many courts as tacit acceptance of the technique). This state of affairs did not last long; numerous articles soon appeared in the literature. These articles pointed out flaws in Professor Tosi’s experimental design, reinterpreted his data, and pointed out important questions. The next section of this paper will discuss the more relevant criticisms gleaned from the literature.

4. Voiceprint Reservations After Tosi

Stress and Voiceprints

Bolt remarks that:

“Not examined in Tosi’s study are some other factors that can change the sounds a speaker produces, factors that can increase the intraspeaker variability — these in turn can increase the probability of error. For example, it is well known that changes in the psychological state of a talker, induced through emotions or other types of stress can cause substantial changes in the characteristics of his speech sound. In a forensic situation additional emotional factors of this kind tend to be present.”\(^{85}\)

---

82. *Supra*, note 57 at 514
83. *Id.*
84. *Id.*
85. Bolt #II, *supra*, note 78 at 532
The exemplar given by a calm suspect long after a crime may differ markedly from a recording taken during the perpetration of that crime. The effects of stress have been found to be spasmodic and inconsistent in some individuals, which would make transformation routines futile.  

Time Interval

As discussed earlier, the matching of non-contemporary spectrograms was attended by greater error than that of contemporary spectrograms. When the known set of voices was 40 people the error rose from 4.5% to 9.8% as the time interval between successive recordings was increased from ‘‘0’’ time to one month.  

Further study is required to see if longer time intervals have higher rates of error: ‘‘... this defect could be highly significant in the forensic situation which may involve samples made years apart.’’

The time interval in State v. Andretta was five years.

Aging

The aural effect of aging of an individual must be taken into consideration when long time intervals are encountered. A West German study found that:

Neither the formant structures of vowels and vowel-like sounds nor the fundamental frequencies determined from spoken sentences consisting of several parts are independent of age. On the contrary, it has been shown that with increasing age the point of concentration of the formants move towards lower frequencies. Moreover, the ability of controlling the pitch frequencies begins to decrease with increasing age. This allows the conclusion that the human phonation system may change predictably with increasing age.

That study examined samples taken from one to ten years after the initial recording. Tosi’s results showed significant increased error after one month.

There is a strong counter-argument to the foregoing criticisms. Greene explains it as follows:

86. Supra, note 6 at 526
87. Bolt #2, supra, note 78 at 532
88. Supra, note 29 at 274-75
To the extent it is assumed that these variables might result in changes in spectrograms produced by the same speaker at different times, the hypothesis usually adopted by voiceprint critics, it is clear that such a consequence would have little or no relevance to increasing false identifications of innocent persons, the spectre which is feared by the same critics. Instead, the result would be the false elimination of actual speaker-suspects because of changes in their voiceprints and a resulting inability of voiceprint examiners to testify to matches on the basis of spectrographic examinations where matches in fact exist.91

Some other noted shortcomings of Tosi's work and the technique generally are:

Telephone Recording

Telephone recording and forensic voice spectrography are almost inseparable. The fidelity of recording equipment used in actual forensic conditions would probably not be comparable with Professor Tosi's laboratory equipment. The random cross-chatter heard on a telephone line will not be the same on the sample recording and the subsequent exemplar. These problems have not been investigated as yet.92

Sample Size

Barnett cogently describes this problem as follows:

Although one would believe that under field conditions an examiner would probably be working with an immense range of possible matches, Dr. Tosi minimizes this variable by suggesting that the "catalogue of known voices is open, true, but limited to a few suspected persons." However, an increase in known set from ten to forty possible matches doubled the probability of making a false identification. What effect an even larger speaker sampling would have is still unknown.93

Professor Tosi suggested that certain factors encountered in forensic situations (but not in his experiments) would improve this situation, (see "Tosi's Rationale" supra) but that remains specualtion until shown by experiments.

Mimicry and Disguise

Professor Tosi suggested that further research should be carried out

91. Supra, note 61 at 194-95
92. Supra, note 7 at 318
93. Supra, note 59 at 721
on disguised and female voices.\textsuperscript{94} The above-mentioned West German Study (see "Aging" \textit{supra}) found that people could successfully disguise their voices so that voiceprint identification was thwarted. W. R. Jones states:

With regard to mimicking, Dr. A. J. Fourcin has conducted experiments which have led him to state that it is possible for a person to imitate another's speech so well that their spectrographic patterns may be so confused that they cannot be distinguished.\textsuperscript{95}

Female Voices

No experimental work has been done on female voices. Researchers cannot extrapolate experimental findings gleaned from male voices, to female voices. Feminine voices have a higher pitch, and would produce different voice spectrographs. We cannot assume that the machine works as well with higher pitch recordings.

Decision Criteria

The most telling criticism is the lack of accepted decision criteria. Bolt, commenting on the Tosi study, says that:

\begin{quote}
The present level of knowledge about personal voice characteristics, their recognition, and how they change under different conditions is still rudimentary. The recent work on speaker identification from spectrograms does not provide any new understanding as to which spectrographic features correlate most clearly or efficiently with the speaker's identity. . . . At the present time . . . the spectrographic identification of a voice by a trained observer appears to rely on a broad assessment of loosely defined points of similarity rather than on a carefully specified set of objectively defined spectrographic attributes. The Tosi experiments, in fact, show considerable disagreement among different panels of observers as to what constitutes a match when they are given the same matching task.\textsuperscript{96}
\end{quote}

The subjectivity of the method is further shown, by Barnett when he notes that:

\begin{quote}
The difficulties inherent in the comparison process of speech spectrograms has led Mr. Fausto Poza of the Stanford Research Institute to claim:

To the present time, no one has been able to describe the nature
\end{quote}

\textsuperscript{94} Supra, note 73 at 2041
\textsuperscript{95} Supra, note 7 at 318
\textsuperscript{96} Bolt #2, \textit{supra}, note 78 at 533
of these characteristics objectively. Nor has anyone been able to devise a systematic method of discounting unimportant dissimilarities that commonly occur in comparing different repetitions of the same speech sounds by one person. Similarly, it has been impossible to quantify the degree to which two sets of spectrograms may resemble each other due to common linguistic content and yet be made different talkers.97

A scientifically valid set of decision criteria is required before objective, repeatable, results can be obtained.

Additionally, a set of standards for purposes of identification is essential to a uniform program for training voiceprint examiners.98

No study has examined the worth of individual voiceprint parameters.

Voiceprint Technician Rarity

One of the peculiarities of the voiceprint technique, which fosters suspicion and mistrust is the lack of trained technicians. The same seven or eight men appear in nearly all of the proceedings in which the admission of voiceprint evidence is sought. Thus, it is their testimony and their testimony alone upon which some fifty courts have based decisions on admissibility.99

Indeed, most of the courts heard the testimony of one expert, Lieutenant E. Nash of the Michigan State Police!

IV. Voiceprints in Court

1. Admissibility Standards

(a) Frye Test

Scientific evidence will be admitted

... if the method used is reliable and if its introduction outweighs the dangers of jury confusion, repetition, unfair prejudice, undue weight and credibility, and the possibility of mistake.100

This discussion will now examine the criteria required to show that a new scientific technique is reliable, and is acceptable

97. Supra, note 59 at 720
98. Id. at 721
99. Id. at 703
scientific evidence. The most widely used test is the one enunciated in Frye v. United States. That test requires that a new method be shown to be generally accepted as reliable within its particular scientific field. It was more eloquently stated in the Frye decision as follows:

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.

Problem #1
This test is the most stringent of those used by the courts; it requires a high level of acceptance from the experts in that particular field. The exact percentage of "converted" experts required was not made clear. Luchetti has stated that:

Perhaps the basic reason for hostility toward the Frye rule is that it requires something greater than a sole expert's acceptance but something less than full acceptance by all experts in the particular discipline. Critics argue that this formulation lacks specificity.

In practice, the courts have decided the "number of experts" question on a case by case basis.

Problem #2
Opponents of the Frye standard claim that it unduly retards use of new scientific techniques.

Frye has also been attacked because it does not allow for the admissibility of a technique which, although scientifically accurate, is yet so new and unique as to be impossible of general acceptance. A reason for this phenomenon is that the development of intra-discipline communication has not kept pace with the development of specialization in scientific research.

Proponents reply that:

This argument, applied to voiceprint evidence, however, assumes that the technique will prove to be scientifically valid

101. (1923), 293 F. 1013 at 1014 (Col. Dist C. A.)
102. Supra, note 100 at 262
103. Id.
and suggests that postponement of admission of voiceprint evidence may be harmful to the judicial process. But what also should be considered is the harm that might be done to the judicial process if, after being accepted, the technique is proven to be unreliable.\(^{104}\)

(b) \textit{Frye Test Modified}

Some courts have used a less stringent test for the admission of a new scientific technique as evidence in trials. For example, in \textit{Commonwealth v. Lykus},\(^{105}\) the \textit{Frye} requirement of "general acceptance in the particular field" was changed to "generally accepted by those who would be expected to be familiar with its use".\(^{106}\)

In effect, this modification shifts the burden of ruling on the acceptability of a technique to the court. It is inevitable that the court will receive evidence for and against the procedure and that court will have to weigh the evidence and come to a decision. This approach presents new problems; some of these are:

1) Jury inability to adequately evaluate technical scientific arguments.

The reliability of the voiceprint technique is left to the factfinder. The argument that a jury of laymen is ill-equipped to handle such an issue recognizes that they can only handle such an issue recognizes that they can only base their determination on their experience and ability to reason from it. Expert testimony is a means of relating unknown experience to the jury.\(^{107}\) At least, a determination of "general acceptance" is closer to the realm of judicial ability than a determination of the scientific validity of the voiceprint technique.\(^{108}\)

2) The modified \textit{Frye} test assumes that the body of "those who would be expected to be familiar with its use" is a reasonable size. There are very few voiceprint experts (as was discussed previously) and those few are all proponents of the technique. Courts have pointed out that:

\[\ldots\textit{Frye} \text{ ensures the existence of a "minimal reserve of experts" who can be called by a party to rebut the testimony}\]

\(^{104}\) Supra, note 57 at 528
\(^{105}\) (1975), 327 N.E. 2d 671 at 677
\(^{106}\) Supra, note 61 at 195
\(^{107}\) Supra, note 57 at 526
\(^{108}\) Id. at 529
of the offering party's experts by attacking the instant
application of the technique. By requiring general acceptance
by the relevant scientific community, *Frye* implicitly
requires that that community exist. 109

A defendant has a right to produce experts to dispute the way in
which the relevant spectrographs were made but no one has ever
challenged an identification itself because there is no one to so
testify. An important . . . avenue for contesting voiceprint
evidence is thereby foreclosed, forcing defendants to challenge
an identification by the more circuitous and less effective
approach of challenging the technique itself. 110 Courts should
not be forced to rely on the opinion of only one examiner.

3) The acceptance of the test in *Frye* should result in greater
uniformity in decisions; it is quite unlikely that judges and juries
throughout the country will have the same view about a
particular technique's reliability; it is far more likely that
different judges can reach the same result on the question of
general acceptance in the scientific community. 111

It is interesting to note that most decisions that have refused
to admit voiceprint evidence have used the *Frye* test and those
that did admit voice spectrograms used a modified *Frye* test.

2. Analysis of Appellate Voiceprint Cases in the United States

Before publication of the Tosi study in 1971, the only court of
appeal in the United States to uphold the admissibility of voiceprint
evidence was the United States Court of Military Appeals in the
case of *United States v. Wright*. 112 In *Wright* the admissibility
question was governed largely by a permissive military court rule
and the majority ignored the general acceptance test as put forward
in *Frye*. 113 According to Barnett

Kersta's testimony was admitted on the basis that he qualified as
an expert at what he did, and not that what he did was reliable (to
the satisfaction of the scientific community). 114

In the years between 1967 and 1971 the issue of voiceprint
admissibility was considered only twice by appellate courts in the

109. *Supra*, note 29 at 290
110. *Id.*
111. *Id.*
113. *Supra*, note 59 at 708
114. *Id.*
United States: *People v. King;*\(^{115}\) *State v. Cary.*\(^{116}\) The courts in both these cases rejected the admissibility of voiceprints.\(^{117}\) However, Barnett points out that in

light of later developments, *Cary* and *King* must be viewed more as an indictment of Kersta and his research than of the voiceprint method of speaker identification.\(^{118}\)

The Tosi study was published in 1971 and since that time appellate courts in five states and three federal circuits have considered the matter, with the overwhelming majority favouring admissibility.\(^{119}\)

However, it would be incorrect to jump to any general conclusions based on any such simple-minded statistical analysis of the cases. What follows is a very brief analysis of the recent appellate cases in those jurisdictions which have come down in favour of voiceprint evidence.

The State Courts

a) *Minnesota* — The case of *State ex rel Trimble v. Hedman,*\(^{120}\) was an appeal from the dismissal of a habeas corpus petition. The issue there was whether the issuance of an arrest and a search warrant could be justified on the basis of voiceprint analysis alone. The issue was limited to the sufficiency of evidence to show probable cause and the court was not concerned with sufficiency of proof to sustain a conviction.\(^{121}\)

b) *New Jersey* — The case of *State v. Andretta*\(^{122}\) was an appeal from a denial of the prosecution’s request for an order compelling the defendant to produce voice exemplars. The New Jersey Supreme Court allowed the appeal citing developments in the voiceprint field since their previous decision in *Cary.* They did not, however, decide that voiceprint analysis would be routinely admissible at trial and as a matter of fact they directed that if the prosecution desired to introduce evidence of identification by voiceprint analysis at trial there would first

\(^{115}\) (1968), 266 Cal. App. 2d 437  
\(^{116}\) (1970), 56 N. J. 16  
\(^{117}\) Supra, note 6 at 513  
\(^{118}\) Supra, note 59 at 711  
\(^{119}\) Supra, note 6 at 513  
\(^{120}\) (1972), 192 N.W.2d 432  
\(^{121}\) Supra, note 6 at 514  
\(^{122}\) Supra, note 89
have to be a pretrial hearing to determine whether the voiceprint technique of speaker identification was sufficiently reliable to warrant admission of such evidence.\textsuperscript{123}

c) Florida — In both of the two intermediate appellate decisions in Florida, \textit{Worley v. State},\textsuperscript{124} and \textit{Alea v. State},\textsuperscript{125} it was held that spectrographic analysis was properly admissible to corroborate identification by other means.\textsuperscript{126}

d) California — Two of three intermediate appellate decisions handed down in California since 1971 have upheld the admissibility of voiceprint analysis.\textsuperscript{127} However, in one author's opinion\textsuperscript{128} there was a misapplication of a modified \textit{Frye} test in the case of \textit{Hodo v. Superior Court},\textsuperscript{129} and in the case of \textit{People v. Kelley},\textsuperscript{130} the court placed emphasis on the fact that the defense had not offered any experts of literature on its behalf.\textsuperscript{131}

e) Massachusetts — In two recent cases the Supreme Judicial Court of Massachusetts approved of identification evidence based on the voiceprint technique: \textit{Commonwealth v. Lykus};\textsuperscript{132} \textit{Commonwealth v. Vitello}.\textsuperscript{133} In \textit{Lykus} the court relied on a modified \textit{Frye} test in approving of the technique\textsuperscript{134} and saw fit to add a cautionary note to its support of the technique:

\begin{quote}
We add that the admission of expert testimony as to spectrographic analysis should be subject to the closest of judicial scrutiny, particularly in any case where there is an absence of evidence of voice identification other than that of the voiceprint or where, but for the voiceprint, there would be insufficient evidence to warrant any inference of the defendant's guilt.\textsuperscript{135}
\end{quote}

\begin{footnotes}
\textsuperscript{123} Supra, note 6 at 515
\textsuperscript{124} (1972), 263 So.2d 613 (Fla. App)
\textsuperscript{125} (1972), 265 So.2d 96 (Fla. App.)
\textsuperscript{126} Supra, note 6 at 515
\textsuperscript{127} Id.
\textsuperscript{128} Supra, note 59 at 717
\textsuperscript{129} (1973), 30 Cal. App.3d 780
\textsuperscript{130} (1975), 49 Cal. App.3d 214
\textsuperscript{131} Supra, note 6 at 516
\textsuperscript{132} (1975), 327 N.E.2d 671
\textsuperscript{133} (1975), 327 N.E.2d 819
\textsuperscript{134} Supra, note 6 at 513
\textsuperscript{135} \textit{Commonwealth v. Lykus} (1975), 327 N.E.2d 671 at 679.
\end{footnotes}
The Federal Circuit Courts

The issue of admissibility has arisen before three circuits of the United States courts of appeal. In *United States v. Addison*, the admission of voiceprint evidence at trial was held to be error (non-prejudicial) since, relying in part on the 2nd Bolt study, the court said the technique did not meet the requirements of the *Frye* test. The other two circuits to consider the issue ruled in favour of admission: *United States v. Baller* and *United States v. Franks*. However, in neither of these two cases did the defense present any expert witnesses to rebut the government’s claims in regards the reliability of the technique.

Thus, while it is true that most of the recent appellate decisions in the United States have upheld the use of spectrography as forensic evidence, as Siegal says:

In almost all of these decisions, however, a general sense of hesitancy was readily apparent, although manifested in various ways... admission of such evidence should be subject to closest scrutiny... introduction of voice spectrograms was allowed for corroboration purposes only... decided only that spectrographic evidence was sufficient to show probable cause for the issuance of warrants... it was pointed out that no expert testimony or literature had been offered by the defense. This is certainly something less than a ringing endorsement of this technique by the courts.

It is also impossible to overlook the fact that in every case discussed, Oscar Tosi and Ernest Nash have been the key — and sometimes the only — expert witnesses... The amount of weight and credibility given by the courts to the testimony of these two men and to the Tosi study cannot be overemphasized.

Barnett also points out that for quite some time after the Tosi study the courts in the U.S. seem to have considered the issue of scientific reliability as having been resolved by the Tosi study. However, as noted above there still remains some reservations in the scientific community which are perhaps not given proper emphasis in court because of the dearth of experts available to put forward a view contrary to that of Tosi.

---

137. (1975), 519 F.2d 463 (4th Cir.)
138. (1975), 511 F.2d 25 (6th Cir.)
139. *Supra*, note 6 at 516
140. *Id.* at 517
141. *Supra*, note 59 at 714
3. The Canadian Cases

To the best of the authors' knowledge the issue of voiceprint admissibility has arisen in only two Canadian cases to date: Regina v. Montani;142 The Queen v. Medvedew, an as yet unreported trial heard before Mr. Justice Wilson and a jury in the Queen's Bench Assize Court, Brandon, Manitoba, March 22 to 26, 1976.143 This latter case is presently on appeal. In both cases, voiceprint analysis were admitted into evidence, but it should be noted that in each case such analysis was corroborative of aural identification. It should also be noted that Montani was only a preliminary inquiry on the admissibility question and that Medvedew has only been considered at the trial level yet. What follows is a brief criticism of each case.

(a) Montani144

The authors' criticisms with the ruling in this case are threefold:

1. The judge seems to assume that the Tosi study was the end of the reliability question.145 In view of the recent criticisms of the Tosi study this may have been a somewhat premature assumption. It appears that the judge may also have been misinformed on the problem of mimicry.146

2. Nowhere in the decision does the judge discuss the evidentiary standard of admissibility which should be applied to scientific evidence. Instead he appears to be using a statistical analysis of recent American cases to justify his ruling for admissibility. The judge merely lists some thirty or so American cases147 and then simply says:

Many of those cases have been made available to me by Mr. A. Zuraw for the crown and they deal with the question of admissibility and generally speaking have admitted the voiceprint evidence. I therefore rule that Lieutenant Nash's testimony based

143. The authors are extremely grateful to Cpl. K. F. Taylor with the office of the Commissioner of the R.C.M.P., Special “I” Branch, in Ottawa, who, in private communication dated 5 November, 1976, kindly provided us with a copy of the charge to the jury in Medvedew as taken from the transcript of the trial.
144. See also the annotation Spectrographic Analysis: Its Admissibility and Weight (1974), 26 C.R.N.S. 347 following the report of Montani. The annotation was written by Mr. Anton Zuraw who appeared for the Crown in Montani.
145. (1974), 26 C.R.N.S. 345
146. Id.
147. Id.
on spectrographic analysis is admissible as evidence.\textsuperscript{148}

As noted before, such an approach ignores the hesitancy often expressed by the U.S. Courts which have admitted such evidence. As this was the first reported case in Canada dealing with the admissibility of voiceprint analysis it would have been desirable that the judge expanded his analysis of the U.S. case law in giving his ruling.

3. Nowhere in the report does it appear that the defence presented any expert witnesses on their behalf, so that the judge was probably getting only one side of the story.

(b) \textit{Medvedew}

This recent case was concerned with a bomb threat to a high school. The threat was expected and a recording was consequently made, (there had been a similar threat one year before). The investigating officer recognized the voice as well as four other officers and Mr. Medvedew was subsequently arrested. The suspect agreed to give a voice sample, which was made over the phone and this sample was sent to R.C.M.P. Cpl. Ken Taylor in Ottawa. Cpl. Taylor, an apprentice examiner, felt that he could not make a positive identification and so the recording was sent to Det. Sgt. Lonny Smyrkovsky of the Michigan State Police. Mr. Smyrkovsky made a positive voiceprint identification and gave evidence before Mr. Justice Wilson at a preliminary hearing and at trial. The learned judge did not hold a \textit{voir dire}.\textsuperscript{149} Both the crown and defence introduced expert evidence in the field of the human voice. These experts expressed opposing opinions on the suitability of voiceprint analysis in forensic situations.\textsuperscript{150} How much weight the jury placed on the evidence of voiceprint analysis is uncertain, however, since sever al persons testified that they aurally recognized the voice on the bomb threat tape as being that of the accused. The jury were also given a tape with the bomb threat as well as two samples given by the accused on it so that they might make their own aural comparison.\textsuperscript{151} It is of note that the U.S. voiceprint analyst did not bring with him the spectrograms which he had made from the tapes

\textsuperscript{148} Id. at 346
\textsuperscript{149} Personal communication from Mr. L. R. McInnes, Crown Attorney, Manitoba.
\textsuperscript{150} \textit{Medvedew} transcript, p. 409
\textsuperscript{151} Id. at 413
but merely offered his opinion, that based on his analysis the voice in the samples was the same voice which made the bomb threat.\textsuperscript{152}

Mr. Medvedew was convicted and sentenced to 18 months imprisonment. He represented himself in an appeal and that appeal was rejected (as being prolix and irrelevant) and was proposed sine die. Mr. Medvedew is currently working to earn enough money to be able to hire appeal counsel.\textsuperscript{153}

V. \textit{Conclusions and Recommendations}

The authors are of the opinion that voiceprint evidence should not be admissible at this time. The only exception to this which we would be prepared to accept is in cases where the voiceprint evidence is merely corroborative of identification by other means. However, even in such cases there is a danger that the factfinder will place excessive weight on this sort of evidence unless he is confronted with evidence via literature or expert testimony which points out the weaknesses and deficiencies of the technique as it exists today. In any event, the time has not yet arrived when the justice system should feel safe in relying solely on voiceprint analysis in determining the guilt or otherwise of an accused.

The problem is that more experimental research is required to resolve the questions which remain about the technique. The Tosi study was fine as far as it went, but it was not the final answer to the reliability issue. More such work remains to be done, for example, in areas of mimicry and disguise, time lag, female voices, and the recording conditions in the forensic situation. The present state of affairs in the area of decision criteria, which permits subjectivity on the part of examiners, is to the authors quite unsatisfactory. The determination of a set of scientifically valid, objective decision criteria would be a significant improvement. Work is being done in the area of speaker recognition using computers and though much is yet to be done in this area, the method appears to hold great promise.\textsuperscript{154} However, for the present, the authors believe the Canadian courts would be wisest to take a wait and see approach.

\begin{flushleft}
\textsuperscript{152} \textit{Id.} at 406
\textsuperscript{153} \textit{Supra,} not 149
\textsuperscript{154} Michigan State Police, \textit{Voice Identification Research,} P.R. 72-1, February 1972, p. 32
\end{flushleft}