

DECLARATION OF HERBERT JOE, M.A., J.D., LL.M., B.C.F.E.

"I, Herbert Joe, the Affiant, a resident of the State of Utah, being over the age of 18 years and competent to testify, make this affidavit according to my own personal knowledge, information and belief and upon oath, I state, allege and aver, under penalty of perjury pursuant to 28 USC §1746, that the following is true:"

"I am the managing partner of Yonovitz & Joe, L.L.P. I have 4 degrees, including 2 science degrees (B.S., M.A.) and 2 law degrees (J.D., LL.M.). The following are a result of my expertise and experience in the area of forensic audio/video: Board Certified Forensic Audio/Video Examiner; Diplomat, American Board of Forensic Examiners; Diplomat, American Board of Law Enforcement Experts; Licensed Instructor, Texas Board of Private Investigators; Member, Evidence Code Committee of the Oklahoma Bar Association; Former Board Member, Forensic Expert Witness Association; Board of Legal Advisors, American Guild of Court Videographers; Charter Member, Legal Advisory Board, American College of Forensic Examiner Institute; Member, The Forensic Commission; Fellow, American Guild of Court Videographers; and Fellow, American College of Forensic Examiners. I am also a Certified Mediator, licensed to practice law in Texas and Oklahoma, a Registered Patent Attorney with the United States Patent & Trademark Office, a Registered Patent Agent with the Canadian Intellectual Property Office, a Member of the College of the State Bar of Texas, and Adjunct Faculty (Grad. Law Classes) at the University of Phoenix. My 24-year involvement in the area of forensic audio/video includes expert testimony in state and Federal courts in civil and criminal cases throughout the U.S., as well as overseas, giving regional, national and international presentations, authoring peer-reviewed publications and being interviewed or consulted with throughout the news and entertainment industries domestically and foreign. As for the forensic authenticity analyses of audio evidence, I have only been *Daubert* challenged one time. A true and accurate copy of that Order is attached.

"Clearwater, Idaho Attorney Wesley Hoyt has retained our services to review the audio evidence, and related documents, in the past case involving Mr. Edgar Steele, Criminal Case No. 10-CR-148-BLW, In the United States District Court, District of Idaho. After reviewing everything presented to us, as forensic audio experts, to the extent possible, we would have recommended that trial counsel should have done things much differently, as that could have materially affected the outcome. Hence, below is a list (in no particular order) of what should have at least been done, and why, as far as the objective and unbiased forensic analyses of the audio evidence is concerned:

Issue #1: Chain of custody of the original June 09 and June 10 recordings should have been established and verified before trial.

Explanation: This is at least to establish the presumptive integrity of the recordings at issue. This is also to allow all parties involved to know what generation of recordings are at issue to settle the questions whether the version being played during trial for the jury



was the highest generation available, was authentic and whether it was an exact, bit-for-bit accurate copy provided by the government to defendant's attorney and his experts.

Issue #2: A full explanation and objective justification by Agent Sotka and/or his supervisor as to why each of the original ADS recordings were deleted should have been disclosed, established and verified before trial.

Explanation: Agent Sotka's testimony for the deletion of the June 09 recording is not fully and satisfactorily explained or justified. If it was to free up space on the media, then it should have been established before and/or at trial as to the nature of the media, the media's storage capacity, standard and/or written procedures as to the proper handling and preserving of original (audio) evidence, etc. If it were to free up space on the media, Agent Sotka should have been required to explain on the record as to how a 45 minute recording is filling up the media as to its capacity and the amount, if any, of other audio recordings thereon (were these from other cases or from this case) such as to necessitate its deletion.

Even if this were done, why would the June 09 recording, for example - at approx. 45 minutes in duration - require its deletion? In other words, was the media near full prior to the June 09 recording? If so, were the other recordings relevant to the Steele case? If not, why were they not deleted before the informant was sent out with the FLEX8F recorder to free up space? And since the media was freed up before the June 10 recording, then space could not have been a compelling issue for that one 45-minute recording. And if space is an issue, then why were these recordings deleted after being uploaded onto the PC? (Surely, 2 45-minute recordings would not take up much space at all on an endless supply of inexpensive, external hard drives.) Also, we do not have any explanation how, when, why or even if the putative original June 10 recording has been deleted. If so, we do not know why it was deleted.

Furthermore, the single-board FLEX8F recorder, manufactured by Adaptive Digital Systems, Inc. of Newport Beach, CA, has a stereo recording capacity of 13.4 hours. If there were no other audio files on the FLEX8F at issue, then its remarkable to delete original evidence of 45m16s just to save less than 6% of the total available space. On the other hand, if there were files on the FLEX8F recorder at issue, then they were either files that are relevant to the Edgar Steele case that we do not have, or they are irrelevant to the case - which begs the question of why these irrelevant files were not deleted when originally made, if there was, in fact, a written policy of such deletion.

Issue #3: Agent Sotka's capacity and qualifications to properly operate the ADS FLEX8F recorder should have been established and verified before or at trial, as well as the methods by which he made the recordings.

Explanation: This is obvious. From a thoroughness or due diligence standpoint, the basic competency of the operator of the recording equipment should be established not only from its basic operations, but also with respect to the handling and transferring of the original audio evidence. And in hindsight, this information would be very important to

have now, as any forensic examiner would need such information to be accounted for, e.g., operator competency, chain of custody, how audio files were transferred from one media to another, etc.

"Issue #4: Equipments' (e.g., ADS, K1 and K2) integrity should have been established and verified before trial, including but not limited to, any and all repair records.

Explanation: Our understanding is that the ADS recorder referred to as "K1" is the putative original recording system for both the June 9 and the June 10 recordings; although the FBI "Report of Examination" report dated April 13, 2011 states that recorder "K-1" was initially received on April 07, 2011 from Agent Sotka as if it was the device on which the recordings were made. It has been held out also that it, *i.e.*, K1, was not working properly before or upon its arrival at Quantico for analyses. And then it was held out that another recorder, namely ADS recorder "K2", received on April 11, 2011, is actually the putative original recorder for both the recordings at issue. Given this ambiguity, verification of which recorder actually made the recordings at issue should have been undertaken to scientifically determine which of the recorders was used for each day's recording. Regardless, along with competency of the operator, all logs, use records, repair records, etc., of K1 and K2 should have been ascertained prior to trial as an integral part of the forensic analyses of the audio evidence. All of this is quite remarkable in that K1 and K2 have very different and distinct Government Property Numbers (namely, F1840800 and F2378520, respectively).

"Issue #5: Forensic examination of the highest generation recording for each of the putative original recordings should have been made and analyzed before trial, as well as forensic comparisons between all copies made by the government and the highest generation recording available.

Explanation: Although one can authenticate a copy to the extent that it is a copy, the most thorough authenticity analyses is the forensic examination of the putative original recording. The next most thorough authenticity analyses would be of the 1st generation copy, then the 2nd, etc. So, if the original evidence is not available, then forensic examination of the highest generation recording available is necessary. This was not done in this case. Without access to the original recordings to use as a comparative for authenticity, there can be no scientific assurance that the putative first generation copy or any other generation of copies, is or are authentic unless there is verification that the transfer from the ADS device was a "bit-for-bit" accurate transfer of data. Without such proof, authenticity cannot be assured.

"Issue #6: Forensic examination of the putative original recorder(s), e.g., K1 and K2, should have been at least requested, if not granted and analyzed before trial.

Explanation: Having access (whether on-site or not) or possession of the putative original recorder(s) could have (and still could) provided invaluable forensic information about the authenticity of the putative original recordings. For example, non-invasive,

non-destructive test recordings could have established an eletroacoustic baseline in which to compare the recordings at issue, which could ultimately aid in determining the recordings' authenticity, or to establish evidence of falsification.

Issue #7: Determination of, and then forensic examination and comparison of Plaintiff's Exhibits 21 and 22 as compared to Q1 and Q2 with the highest generation recording for each of the putative original recordings should have been made and analyzed before trial.

Explanation: From a basic thoroughness or due diligence standpoint, as well as it being obvious, the court and all involved should know or have confidence that what audio evidence is played at trial, *i.e.*, Trail Exhibits 21 and 22, is not only free from falsification but also a true and accurate copy of the original audio evidence. And, in hindsight, there is testimony (from at least Ms. Cyndi Steele and daughter Kelsey Steele) that what was played at trial is not a true and accurate or reliable copy of another generation copy supplied by the government as the putative original audio evidence.

Issue #8: In addition to (and separate from) the chain of custody issues per putative original recording, every step of transfer of a digital file should have been established, explained in detail and verified before trial.

Explanation: For example, it is very material to know exactly how the putative original recordings were transferred from the ADS recorder to a specified computer; were there options in settings to export those files? What software was used to do the transfer? Was a USB cable or port used for the connection? If not, then what was the connection? Who did the transfers? How did the person making the transfer from the ADS recorder to the computer know that the transferred file was a bit-to-bit accurate, exact digital clone of each of the original ADS FLEX8F audio files? What storage device was used to store the first uploaded audio files? How was this evidence stored? What is the policy for the handling and storage of such audio evidence? Has the PC in which the original ADS files were uploaded ever compromised, repaired, etc.?

Issue #9: The FBI's Forensic AVI unit at Quantico analyzed "Q1" and "Q2" recordings in this case. What (generation) were these 2 recordings relative to the original ADS files uploaded to what computer? This should have been established and verified before trial.

Explanation: This is an integral part of the chain of custody of the original audio evidence. The analyses of Quantico is only as reliable as the evidence that they receive. Is, for example, Q1 a direct archive of the original June 09 recording off of the FLEX8F at issue, or is it a subsequent generation recording? Were any of the copies not bit-to-bit digital clones of the original June 09 recording from the FLEX8F recorder at issue?

Issue #10: What computer was actually used to upload the 2 putative original ADS recordings? This should have been established and verified before trial. Is the

same hard drive still in that computer? Does that hard drive still hold the data from the original uploads? Has the data from the original uploads been transferred to another hard drive? Is that computer in service at the present time? If not, what is the history of that computer and where is the hard drive at the present time? If the data from the original upload was transferred to another hard drive, then defense counsel should have timely asked the Government to designate all places where has the data from the original upload been transferred.

Explanation: This is an integral part of the chain of custody of the original audio recordings. Testimony states that the original audio recordings were "downloaded" (pp.353-354, Agent Sotka trial testimony) before they were deleted (p.360, Agent Sotka trial testimony). Knowing and verifying the integrity of the relevant computers is essential information. It potentially can affect the authenticity of the original recordings, as well as potentially affect the authenticity or integrity of subsequent copies thereof, even in the absence of malfeasance.

Issue #11: Whatever recordings were played at trial, it should have been established and verified before trial exactly what generation recordings these were.

Explanation: It stands to reason that any audio or video evidence played at trial must be a reliable copy of the original audio or video evidence. There is testimony that what was played at trial is materially different than what was heard before trial. Having whatever audio evidence that was to be played at trial forensically analyzed for authenticity as well as thoroughly comparing what was to be played at trial to the (putative) original evidence achieves the goal of assuring that what is played at trial is a true and accurate copy of the original audio evidence. Attached is the Affidavit of Cyndi Steele in this regard, which indicates that a new sound may have been introduced into the recording that she had not heard before, the so-called "tic-tak" noise.

Issue #12: Since "K1" was not working by the time it arrived at Quantico, it should have been established and verified before trial as to the integrity, operability or lack thereof of K1 before and during June 9 and 10, as well as since that date, as well as how often it was used since June 9 and 10 until arriving at Quantico.

Explanation: If the integrity of operability of "K1" was an issue at least when it arrived at Quantico, then that necessarily begs the question of how long prior to its arrival at Quantico did it not operate properly. And if this is the case, how do we know it's integrity or lack thereof on June 10 and June 11, 2010? Having as much information about the K1 recorder could shed light as to the authenticity of the recordings at issue.

Issue #13: It should have been established and verified before trial as to the chain of custody of K1 and K2, as well as any logs or documentation about either of their use, and any all repair or malfunctioning records of both K1 and K2.

Explanation: There appears to be a question as to which recorder(s) made the June 09 and June 10 recordings at issue. K1 and K2 have different Government property

numbers. It's hard to believe that there is no documentation as to which recorder made which recording, or which recorder has repair records or the like.

Issue #14: If each of the putative original recordings were transferred to one or two different computers, then it should have been established, satisfactorily explained and verified before trial as to why those files are no longer on that/those computers. Likewise, if each of the putative first generation recordings were transferred to one or two different computers, then it should have been established, satisfactorily explained and verified before trial as to why those files are no longer on that/those computers.

Explanation: These are additional important aspects of chain of custody. From the records available and reviewed by the undersigned and based on the fact that the FLEX8F recorder can record up to 13.4 hours continuously, it is not clear to the undersigned and the trial counsel should have timely received adequate explanation as to why the original recordings (just 35 and 45 minutes in duration) had to be deleted; and it should be clear as to how these original recordings were made, archived, copied, deleted, etc.

Issue #15: Transcripts of each of the putative original recordings made by or on behalf of the Government should have been scrutinized and verified before trial, and challenged if necessary, or the defense should have proffered their own version.

Explanation: The record apparently reflects that the defense did not challenge the accuracy of the Government's version of the transcripts, and the defense did not proffer their own (transcript) version of the recordings. If applicable, either should have been done.

Issue #16: Certified cell phone records of Mr. Larry Fairfax should have been ascertained, scrutinized and compared to all available acoustic and linguistic information from the audio recordings before trial.

Explanation: It appears that Mr. Fairfax was talking on his cell phone during the June 9 and/or June 10 recordings at issue. Having certified cell phone information to cross-reference the June 9 and/or June 10 recordings can substantiate the timing, for example, for the recordings, or they may substantiate falsification of the recordings. Either way, having additional information to cross-reference the June 9 and 10 recordings at issue can only help authenticate the recordings at issue.

Issue #17: Since there are or may be "unintelligible" portions in the audio evidence, a forensic audio expert should have been retained to digitally enhance those section before trial.

Explanation: The defense should have retained an audio expert to enhance every portion that is inaudible and/or unintelligible. Revelation of words or speech could be very important, regardless of which "side" it may help. Having a very thorough and reliable transcript from audio that has been enhanced when needed is obviously

important from the basic standpoint of knowing what words one is dealing with. Secondly, a thoroughly intelligible and accurate transcript can provide the basis for a forensic linguistic or discourse analysis of the speech and dialogue between individuals.

Issue #18: Since it is alleged that some of the speech attributed to Mr. Edgar Steele in the June 09 and/or June 10 recording is actually not that of Mr. Steele, then trial counsel should have retained the forensic services of a voice identification/elimination expert before trial.

Explanation: Hiring a competent voice expert - unlike the one in *U.S. vs. Angleton*, 269 F.Supp.2d 892 (S.D.TX 2003) - before trial could have provided invaluable information for trial, as well as direct implication about the authenticity of the audio recordings. Attached is a peer-reviewed article published in the *Law Enforcement Executive Forum* about forensic voice identification/elimination.

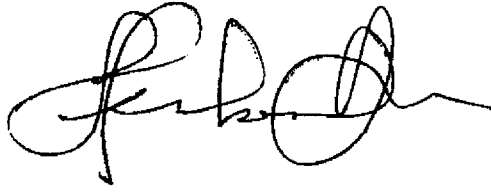
Issue #19: Since forensic discourse or linguistic analyses of the audio evidence always has the potential to provide a legal defense when audio evidence is concerned, then trial counsel should have retained the forensic services of a linguistic expert before trial.

Explanation: Two of the several ways in which objective forensic discourse analyses by a linguistic expert can provide a legal defense includes clarification of understanding or intent amongst the conversants, and secondary authenticity analyses of the audio evidence. Either of these could be very beneficial, and in cases where both are applicable can be a compelling defense.

Issue #20: Information as to what settings Agent Sotka chose and why for the making of the recordings, as well as chosen settings for their transfer should have been ascertained, elaborated upon, scrutinized and verified before trial.

Explanation: Such information is material and relevant to authenticity. The more information an examiner has, the more parameters and characteristics the analyzed audio evidence should have. And not only do settings matter in authenticity, they may matter as far as making copies are concerned. For example, were the audio files compressed or not when made? Generally, compression affects the quality or resolution of the recording. Also, how were the original ADS files archived? What is the chain of custody of the archived files?

"I declare nothing further."



Herbert Joe, M.A., J.D., LL.M., B.C.F.E.

DATE: 08 August 2011

Attachments:

- 1) *Daubert* Order, file stamped Oct. 01, 1999;
- 2) Affidavit of Cyndi Steele; and
- 3) Joe, H., Yonovitz, A. "Legal, Scientific, and Forensic Controversies Over Spectrographic Voice Analysis for Identification or Elimination, Vol. 7, No. 6, Sept. 2007, *Law Enforcement Executive Forum*.



Law Enforcement Executive

FORUM

Law Enforcement Executive Forum • Police Training, Investigations • Vol. 7, No. 6

September 2007

Table of Contents

Editorial	i
Vladimir A. Sergevnin	
Police Training: Investigations	
Making the Collar: Theories of Criminal Investigation	1
DeVere D. Woods, Jr.	
Using Investigative Interviewing Strategies Over the Telephone: Outcomes of an Innovative Training Program	9
Catherine Layton John Nixon Murray Lee	
Paper or Plastic: The Need for Standardized Training and Knowledge Regarding Evidence Collection and Preservation	21
Vicky J. Blair	
A Quick Look at Interrogation	33
Adam Kasanof	
Mandatory Recording of Criminal Interrogations: Issues and Policy Considerations	45
James N. Gilbert	
Legal, Scientific, and Forensic Controversies Over Spectrographic Voice Analysis for Identification or Elimination	51
Herbert Joe Al Yonovitz	
A Police Investigation Meets Forensic Accounting	59
Timothy Pearson Richard A. "Dick" Riley, Jr.	
Your Investigation and the Media	73
Rick Rosenthal	
The Legality of Canine Sniffs and Motor Vehicle Stops Revisited	81
Barry M. Anderson Terry M. Mors	
Our Panoptical Society Where the Invisible Becomes Visible	85
Jill Joline Myers	

Legal, Scientific, and Forensic Controversies Over Spectrographic Voice Analysis for Identification or Elimination¹

Herbert Joe, MA, JD, LL.M., BCFE, Partner, Yonovitz & Joe, LLP
Al Yonovitz, PhD, CCC-A, Partner, Yonovitz & Joe, LLP

Since 1923², the standard for admissibility of scientific evidence in federal courts was governed by the "Frye Test," in which evidence was admitted if it was based on principles "generally accepted" within that scientific community. This basically meant that the relevant scientific community would in effect decide whether a particular scientific methodology would be admitted or not. One significant shortcoming of the Frye Test was its failure to distinguish new or novel scientific or technological procedure. That was the law of the land (at least for federal courts) for 70 years until the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals*³ ruled otherwise in 1993.

The U.S. Supreme Court overruled the Frye test in the *Daubert* case, which is the foundational case for the admissibility of scientific evidence and arguably affects tort reform more than any other single case. The *Daubert* case made the trial judge the "gatekeeper" in admitting the good sciences and barring the junk sciences. The *Daubert* test, now the rule of law for all federal courts, is a nonexclusive list of factors used by the federal courts to determine the reliability of the expert testimony or scientific principles utilized thereof. If an expert relies on unreliable data or methodology, then his or her entire expert opinion is likewise unreliable and should be excluded from the jury.

Some laws have come down since *Daubert* to clarify its rulings: In *General Electric v. Joiner*,⁴ the U.S. Supreme Court held that a trial court's decision to admit or deny expert testimony is reviewed under an abuse of discretion standard at the appellate level. In *Kuhmo Tire v. Carmichael*,⁵ the U.S. Supreme Court held that *Daubert* applied to all forms of expert testimony, not just scientific testimony, and in 2000, Federal Rule 702⁶ was amended to include the logic of *Daubert* and *Kuhmo Tire*.

Currently, admissibility of expert testimony in federal courts is necessarily governed by *Daubert*; however, state courts are not bound by *Daubert*. A significant minority of states (11) have state statutes that specifically adopt *Daubert*⁷ in determining the admissibility of expert testimony in state courts. Thirteen states—including Illinois—are still applying the *Frye* test,⁸ and 11 states, along with the military courts, apply a variation of the *Frye* test.⁹ The remaining states have a history of rejecting both *Frye* and *Daubert*.

Spectrographic voice analysis was developed at Bell Labs in 1941 and introduced forensically in 1961; however, forensic phonetics, which actually predates spectrographic voice analysis, was the topic of a widely published study on

comparing voices.¹⁰ At any rate, the basic premise of spectrographic voice analysis is that each of us has unique vocal characteristics; sinus cavities, vocal chords, and articulators (i.e., lips, teeth, tongue, etc.) confer vocal characteristics, which are individually unique when spectrograms are made. Unfortunately, there is a (mis)perception by the general public that voice ID (VID) is straightforward. This misperception was reinforced by the introduction of the term *voiceprint* or the phrase *voiceprint identification*,¹¹ Hollywood dramas like *CSI*, and the like.

VID or voice spectrographic analysis has been controversial since its forensic outset. This is partially manifested by roughly as many jurisdictions that allowed expert testimony on VID as those that did not before *Daubert*.¹² Since *Daubert*, no federal appellate court has approved the admission of voice spectrographic expert testimony into evidence. The Fifth Circuit federal Court of Appeals has stated that the state of the law concerning expert voice identification is “ambiguous” in the wake of *Daubert*.¹³ One state Supreme Court, however, has allowed expert testimony on VID since *Daubert*, namely, *State v. Coon*, 974 P.2d 386 (AK, 1999) (admission of testimony based on voice spectrographic analysis was not an abuse of discretion).

The Alaska Supreme Court did note that the scientific literature “permit[ted] a conclusion that there is significant disagreement among experts in the field of voice spectrographic analysis regarding the reliability of the technique.”

The same expert who testified in the *Coon* case is the same expert in the *Angleton* case.¹⁴ In *Angleton*, the federal District Court held that the proposed expert testimony on aural spectrographic voice identification method failed to meet the *Daubert* standard for reliability, and the protocol followed by the expert did not protect against several sources of error, further reducing the reliability of the expert’s testimony. Stephen Cain, the expert in the *Angleton* case, testified that he followed the protocol of the American Board of Recorded Evidence (ABRE) in his analysis.

Mr. Cain and Dr. Nakasone testified that the ABRE was formed after a dispute arose among the members of the voice identification board of the International Association of Identification (IAI)¹⁵ over the standards for aural spectrographic analysis. In *Angleton*, Dr. Nakasone testified that the group that left the IAI to form the ABRE felt the IAI’s standards for voice identification were too stringent because they required examiners to obtain a second opinion and to include a statement of accuracy in their reports. Furthermore, . . .

The record before this court shows that the remaining proponents of the use of aural spectrographic voice identification for courtroom testimony are a handful of consultants who apply the techniques for the purpose of litigation. The proponents, including Cain, are not performing scientific research in aural spectrographic voice identification and testifying as experts as an aspect of their research work.¹⁶

In addition, Mr. Cain acknowledged that it is not uncommon for VID analysts reviewing the same recordings to have differing opinions as to the identity of the speakers on the recordings.¹⁷

The state of forensic flux of expert testimony on voice spectrographic analysis is also significantly manifested by the changing views of Dr. Hiroataka Nakasone.

Dr. Nakasone has been working in the field of speech recognition since 1977. While working on his doctorate in speech sciences from Michigan State in 1984, Dr. Nakasone conducted voice spectrographic research for the Los Angeles County Sheriff's Department and on several occasions testified as a proponent of the VID method in courts and administrative tribunals. Since 1992, Nakasone has worked for the Federal Bureau of Investigation (FBI), conducting research in audio forensic identification. His current research is in developing computer-assisted voice identification systems. As late as 1989, Dr. Nakasone testified in *U.S. v. Smith*, 869 F.2d 348, 353-54 (7th Cir. 1989) that the voice spectrographic technique was reliable and had a low error rate.

Recently, Dr. Nakasone has testified that his initial belief that the VID technique is sufficiently reliable for courtroom purposes has eroded over time, as a lack of research efforts have failed to support the underlying premises of the voice identification techniques or to produce reliable testing for error rates. Dr. Nakasone has testified credibly that this failure is the basis of the FBI's approach to voice spectrographic analysis. (The FBI does not permit the use of voice spectrographic analysis for courtroom identification but only for investigation.)

The bottom line, however, is that when the courts ask the wrong questions, they will only get the wrong answers. Likewise, when "experts" primarily or exclusively use a flawed technique, then false positive or false negative answers will occur. As emphasized above, there is a justifiable state of *forensic flux* of expert testimony on voice spectrographic analysis.¹⁸

VID analysis relies extensively on merely matching patterns between spectrograms. Thus, VID analysis (alone) is unreliable because several significant vocal characteristics are not available on spectrograms, and almost every putative expert on VID is not formally (academically or clinically) trained in the speech and hearing sciences¹⁹—a prerequisite to recognizing and distinguishing the vocal characteristics that are not available on spectrograms but are available through other instrumental means including expert aural perceptual analysis.

Specifically, the Aural-Acoustical methodology for Voice/Speaker Identification or Elimination incorporates a single dimensional scaling of the conclusion along a continuum. (Although the Aural-Acoustical method does not rely on spectrographic analysis as its principal basis, it can be supplemented by spectrographic analysis.) This continuum holds as its basis that at one end of the scale, a very high probability (or positive) "identification (match)" exists, and at the other end of the scale, a very high probability (or positive) "elimination (non-match)" exists.

Assuming samples of the recorded evidence and the exemplar recording contains sufficient and intelligible speech materials, the Aural-Acoustical method of speaker identification or elimination has as its basis both acoustical (objective) and aural (subjective) procedures. It is therefore possible that error exists; however, the probability of error decreases with the competency of the forensic scientist. The methodologies have been long established in the speech, hearing, and language sciences and represent ordered analyses of obtained data.

The speaker identification or elimination procedure employed is one in which an unknown voice is taken from an evidence tape and compared to exemplars of a

known voice. In this manner, samples of a number of comparisons between the unknown and known combinations are placed in pairs or composites for direct and repeated comparisons. The Aural-Acoustic method of analysis follows the protocol and standards described in publications as well as a number of presentations to professional organizations, including the Acoustical Society of America and the American Speech, Hearing and Language Association. The principles of this protocol are to provide a basis for voice/speaker identification or elimination that is consistent with the known principles and evidence-based practice of the hearing, speech, and language sciences.

The Aural-Acoustic method has evolved from earlier standards developed by the IAI²⁰ (and the ABRE whose standards practically mirror those of the IAI). For example, section VII.B.5 of the 1996 "Voice Comparison Standards" of the Voice Identification and Acoustic Analysis Subcommittee (VIAAS) of the IAI and section 7.2.5 of the "Voice Comparison Standards" of the ABRE are entitled "Speech Characteristics." Speech and hearing scientists and phoneticians are particularly skilled in forensically assessing speech characteristics.²¹ Examiners trained in spectrogram pattern matching receive little or no training in the assessment of speech characteristics.

Section VII.B.5.i ("Vocal Quality") of the 1996 "Voice Comparison Standards" of the Voice Identification and Acoustic Analysis Subcommittee (VIAAS) of the IAI states . . .

Vocal quality is the perception of the complex, dynamic interplay of laryngeal voicing (pitch, intonation, and stress), articulator movement, and oral cavity resonances. *Since each individual's voice is relatively unique in vocal quality, comparisons can provide important information regarding similarities and differences between the voice samples.* (emphasis added, quoted verbatim by the ABRE in its "Voice Comparison Standards," Section 7.2.5.i. "Vocal Quality"²²)

This Aural-Acoustical method uses a number of instrumental or digital signal processing procedures that delineate the microstructure of various vocal qualities or characteristics, such as those described in the acoustic process below. It utilizes, with due caution, the use of these measures, not to overextend the conclusions that may be offered. Two publications discussing the Aural-Perceptual methods at length are Hollien's *Acoustics of Crime* (Plenum, 1990) and Hollien and Hollien's *Forensic Voice Identification* (Academic Press, 2001). The very significant vocal qualities or characteristics present in speech but not in simple spectrograms include the following:

- Complex co-articulation patterns of vowels and consonants
- Voice quality²³ (e.g., measurements of resonance, vocal fry, and/or nasality)
- Linguistic and paralinguistic features (e.g., prosody, rate, and/or melodic patterns)
- Speech abnormalities (e.g., misarticulation and/or fluency)
- Dialect
- Fundamental frequency²⁴—absolute and variable
- Jitter²⁵
- Shimmer²⁶

In summary, there has been and will always be a great need for various applications of speech processing, speech sciences, and speech technology. Voice or speaker identification or elimination is one of those applications. The forensic comparison of voices or speakers for identification or elimination purposes—by any means—is not infallible. VID, especially by those untrained academically and clinically in the speech and hearing sciences, phonetics, linguistics, etc. is not a reliable means to identify or eliminate voices or speakers.²⁷ There are several significant vocal characteristics that may assist in identifying or eliminating voices or speakers that are not available on spectrograms. The Aural-Acoustic method, with or without supplementing with spectrograms, provides a basis for voice/speaker identification or elimination that is consistent with known principles of the speech, hearing, and language sciences. Under the proper conditions via an academically, clinically, and forensically competent scientist, voice/speaker identification or elimination may be reliably made and academically and forensically tenable between two vocal samples.

Endnotes

- ¹ Copyright is not claimed as to any part of the original work prepared by any government entity. The information you obtain in this article is not, nor is it intended to be, legal advice. You should consult an attorney for individual advice regarding your own situation. Except as stated below, none of the material or derivative works thereof may be reproduced, distributed, republished, downloaded, displayed, posted, transmitted, or copied in any form or by any means, without the prior written permission of the copyright owner. Permission is granted to display, copy, distribute, and download the materials in this article solely for personal, noncommercial use, provided that you make no modifications to the materials and that all copyright and other proprietary notices contained in the materials are fully retained. Any unauthorized use of any material contained in this article may violate domestic and/or international copyright laws, trademark laws, the laws of privacy and publicity, and communications regulations and statutes.
- ² *Frye v. U.S.*, 293 F. 1013 (1923).
- ³ *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 113 S.Ct. 2786 (1993).
- ⁴ *General Electric Company v. Joiner*, 522 U.S. 136 (1997).
- ⁵ *Kuhmo Tire v. Carmichael*, 119 S.Ct. 1167 (1999).
- ⁶ Federal Rules of Evidence Rule 702 (2000) now states that if scientific, technical, or other specialized knowledge will assist the trier of fact in understanding the evidence or help determine a fact issue, then the witness, by virtue of his or her knowledge, skill, training, education, or experience can testify, in the form of an opinion or otherwise, if (i) the testimony is based on sufficient facts or data, (ii) the witness uses scientific methods that are reliable, and (iii) the witness properly applied those reliable scientific methods to the facts of the case.
- ⁷ Connecticut, Indiana, Kentucky, Louisiana, Massachusetts, Missouri, New Mexico, Oklahoma, South Dakota, Texas, and West Virginia

- ⁸ Alaska, Arizona, California, Colorado, Florida, Illinois, Kansas, Maryland, Michigan, Nebraska, New York, Pennsylvania, and Washington
- ⁹ Arkansas, Delaware, Georgia, Iowa, Minnesota, Montana, North Carolina, Oregon, Utah, Vermont, and Wyoming
- ¹⁰ McGehee, F. (1937). The reliability of the identification of the human voice. *Journal of General Psychology*, 17, 249-271.
- ¹¹ Kersta wrote a 1962 article published in *Nature* using those words. See Kersta, L. G. (1962). Voiceprint identification. *Nature*, 196, 1253-1257.
- ¹² See generally *United States v. Leon*, 966 F.2d 1455 (6th Cir. 1992) (admissible); *United States v. Smith*, 869 F.2d 348 (7th Cir. 1989) (admissible); *United States v. Williams*, 583 F.2d 1194 (2d Cir. 1978) (admissible); *United States v. McDaniel*, 538 F.2d 408 (D.C. Cir. 1976) (inadmissible); *United States v. Jenkins*, 525 F.2d 819 (6th Cir. 1975) (admissible); *United States v. Baller*, 519 F.2d 463 (4th Cir. 1975) (admissible); *United States v. Franks*, 511 F.2d 25 (6th Cir. 1975) (admissible); *United States v. Addison*, 498 F.2d 741 (D.C. Cir. 1974) (inadmissible); *United States v. Maivia*, 728 F. Supp. 1471 (D. HI. 1990) (admissible); *United States v. Williams*, 443 F. Supp. 269 (S.D.N.Y. 1977) (admissible); *United States v. Sample*, 378 F. Supp. 44 (E.D.P.A. 1974) (admissible); *United States v. Raymond*, 337 F. Supp. 641 (D.D.C. 1972) (admissible); *United States v. Wright*, 37 C.M.R. 447 (1967) (admissible); *State v. Gortarez*, 686 P.2d 1224 (AZ 1984) (inadmissible); *People v. Kelly*, 549 P.2d 1240 (CA 1976) (inadmissible); *People v. Law*, 114 Cal. Rptr. 708 (Cal. 1974) (inadmissible); *Hodo v. Superior Court*, 106 Cal. Rptr. 547 (Cal. 1973) (admissible); *People v. King*, 72 Cal. Rptr. 478 (Cal. 1968) (inadmissible); *People v. Drake*, 748 P.2d 1237 (CO 1988) (inadmissible); *Brown v. United States*, 384 A.2d 647 (D.C. 1978) (neither); *Alea v. State*, 265 So. 2d 96 (Fla. App. 1972) (admissible); *Worley v. State*, 263 So. 2d 613 (Fla. App. 1971) (admissible); *Cornett v. State*, 450 N.E.2d 498 (IN 1983) (inadmissible); *State v. Free*, 493 So. 2d 781 (LA 1986) (inadmissible); *State v. Williams*, 388 A.2d 500 (ME 1978) (admissible); *Reed v. State*, 391 A.2d 364 (MD 1978) (inadmissible); *Commonwealth v. Lykus*, 327 N.E.2d 671 (MA 1975) (admissible); *People v. Tobey*, 257 N.W.2d 537 (MI 1977) (inadmissible); *State ex rel. Trimble v. Hedman*, 192 N.W.2d 432 (MN 1971) (inadmissible); *Windmere, Inc. v. International Insurance Company*, 522 A.2d 405 (NJ 1987) (inadmissible); *D'Arc v. D'Arc*, 385 A.2d 278 (N.J. Super. Ct. 1978) (inadmissible); *State v. Cary*, 239 A.2d 680 (N.J. Super. Ct. 1970) (inadmissible); *People v. Bein*, 453 N.Y.S.2d 343 (N.Y. Sup. Ct. 1982) (admissible); *People v. Collins*, 405 N.Y.S.2d 365 (N.Y. Sup. Ct. 1978) (inadmissible); *People v. Rogers*, 385 N.Y.S.2d 228 (N.Y. Sup. Ct. 1976) (admissible); *State v. Williams*, 446 N.E.2d 444 (OH 1983) (admissible); *State v. Olderman*, 336 N.E.2d 442 (OH 1975) (admissible); *Commonwealth v. Topa*, 369 A.2d 1277 (PA 1977) (inadmissible); *State v. Wheeler*, 496 A.2d 1382 (RI 1985) (admissible)
- ¹³ See *U.S. v. Drones*, 218 F.3d 496, 503 (5th Cir.2000). Like the *Coon* court, the *Drones* court noted the "uncertainty of the current state of the law regarding the reliability and admissibility of expert voice identification evidence" (218 F.3d at 504).
- ¹⁴ *U.S. vs. Angleton*, 269 F.Supp.2d 892 (S.D.TX 2003).
- ¹⁵ One of the authors, Dr. Yonovitz, is a former member of the certification and standards committee of the International Association of Identification (IAI).

- ¹⁶ *Id.* at 902.
- ¹⁷ *Id.* at 904, citing Docket Entry No. 160, p. 158, 1.10-1.14.
- ¹⁸ One of the authors, Al Yonovitz, is an associate professor of the speech and hearing sciences with a doctorate in acoustics and having been involved in the academic, scientific, clinical, research, and forensic aspects of vocal production for over 30 years, opines that voice/speech/speaker identification via the Aural-Acoustic method, with or without secondary analyses via spectrography, is an accepted academic method for VID.
- ¹⁹ The authors recognize Dr. H. Nakasone, Dr. Harry Hollein, and Dr. Tito Poza as peer experts in the forensic area of the speech sciences. There may be other U.S. expert or forensic examiners qualified to perform proper VID via the Aural-Acoustic method.
- ²⁰ The IAI ceased certifying voice identification examiners in 1999 and ceased all voice identification activity in December 2002.
- ²¹ One of the authors, Dr. Al Yonovitz, has written numerous publications, given international presentations, produced abstracts, and taught undergraduate and graduate courses related to speech and hearing.
- ²² Approved by the ABRE Voice ID Board in April 1999.
- ²³ Voice quality encompasses the perception of the listener of the overall sound of the talker's voice. Just as different musical instruments produce different wave compositions, the human voice is similar. It is this overtone structure or timbre that can differentiate one voice from another.
- ²⁴ The perceived pitch is the psychophysical correlate of fundamental frequency.
- ²⁵ Jitter is a frequency perturbation of the glottal source signal.
- ²⁶ Shimmer is amplitude perturbation of the glottal source signal.
- ²⁷ Spectrograms may be very useful to engineers and as a voice analysis (e.g., speech pathology) tool.

Herbert Joe is an attorney who has four degrees, including two science degrees and two law degrees. The following are a result of his expertise in the area of forensic audio/video: Board Certified Forensic Audio/Video Examiner; Diplomat, American Board of Forensic Examiners; Diplomat, American Board of Law Enforcement Experts; Licensed Instructor, Texas Board of Private Investigators; Board of Legal Advisors, American Guild of Court Videographers; Fellow, American Guild of Court Videographers; and Fellow, American College of Forensic Examiners. His 21-year involvement in the area of forensic audio/video includes regional, national, and international presentations. He has also authored publications and been interviewed by ABC, BBC, and FOX television. Joe is a partner in Yonovitz & Joe, LLP, a team of forensic audio/video analysts, experts, and consultants.

Al Yonovitz, PhD, has a doctorate in physiological and psychological acoustics and 30+ years of teaching and research, including appointments at the Speech and Hearing Institute; Graduate School of Biomedical Sciences; the School of Public Health at the University of Texas Health Science Center, Houston; the Baylor College of Medicine; the Department of Biomedical Engineering at the University of Houston; and the Conley Speech and Hearing Center, University of Maine. He has been a consultant to the VA Hospital, Houston, on voice research in psychiatric patients. Undergraduate and graduate classes taught by Professor Yonovitz include Auditory Systems and Disorders, Audiology, Seminar in Fluency Disorders, Biomedical Instrumentation, Industrial Audiometry and Hearing Conversation, Computer Applications in Speech Pathology and Audiology, Hearing and Speech Science, Research Methods in Speech Pathology and Audiology, Physiological and Psychological Acoustics, Special Topics: Middle Ear Mechanics, Special Topics: Audiometry with the Difficult to Test, Special Topics: Measurement of Voice, Hearing Impairment, Anatomy of the Speech and Hearing Mechanism, Aural Rehabilitation, Speech Science, Introduction to Audiology, and Audition. He is a partner in Yonovitz & Joe, LLP.

Wesley W. Hoyt, ISB #4950
Attorney at Law
165 Deer Field Dr.
Clearwater, ID 83552
Ph: 208-983-0212
Fax: 208-926-7553
email: hoytlaw@hotmail.com

UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF IDAHO

UNITED STATES OF AMERICA,
Plaintiff,

v.

Criminal Case No. 10-CR-148-BLW

EDGAR STEELE,
Defendant.

AFFIDAVIT OF CYNDI G. STEELE

State of Idaho)
 ss.
County of Bonner)

I, Cyndi G. Steele, the Affiant, a resident of the State of Idaho, being over the age of 18 years and competent to testify, according to my own personal knowledge, information and belief and upon oath, state, allege and aver under penalty of perjury pursuant to 28 USC §1746 as follows:

1. I am the wife of Defendant Edgar J. Steele and was designated as one of the two victims of a murder-for-hire plot supposedly developed by my husband in 2010. I attended my husband's trial in Boise, Idaho from April 26 until the jury reached a verdict on May 5, 2011. The verdict is what I consider to be a wrongful conviction of my husband on all four felony counts of the indictment.

2. Before the trial and on March 8, 9 and 10, 2011, I listened repeatedly and intently to the recordings that the Government said it would be offering in evidence against my husband. These were supposedly recordings of conversations between my husband and Larry Fairfax, our handyman, that took place on my property in a horse barn. Fairfax was acting as the FBI's confidential source ("CS") or confidential informant who supposedly engaged my husband in conversation on June 9th and 10th, 2010 so that such conversation could be recorded. Both recordings supposedly started at 6:02 p.m. and lasted over half an hour each. The CS, Mr. Fairfax, purportedly carried a recording device in one of his pockets that supposedly picked up various sounds, including what was represented to be the sound of Mr. Fairfax' voice and the sound of my husband's voice, on both June 9th and 10th, 2010. It was during these conversations that my husband supposedly was plotting to murder me and my mother.

3. During the trial, the Government took the position that, no one else, other than the CS and my husband, was present when the conversations took place and the recordings were made without the input from any other person. The Government also took the position that the recordings were the uncut and unedited continuous audio recordings, each made without interruption from the time the recorder was turned on by FBI Agent Sotka to the time that it was turned off by him.

4. The Government also represented that the recordings I listened to on March 8, 9 and 10, 2011 were true and accurate copies of the original that was made on the device carried by the CS on June 9th and 10th when they supplied those copies to my husband's attorney.

5. On March 8, 9 and 10, 2011, I gathered with several other people for the express purpose of listening to these recordings in order to determine what was being said and to see if the voice ascribed to my husband was in fact his voice. I listened to both recordings numerous times and compared what I was hearing to a transcript of the conversations which indicated "inaudible" in many places. By careful repeated listening I was able to discern many of the words that had been designated as "inaudible" on the transcript.

6. When I was in court as an observer on April 27, 2011, I listened to two recordings played by the Government represented to be the recordings of June 9th and 10th, presented by the Government to be the same, uncut, unedited recordings that I had listened to on March 8, 9 and 10, 2011. While these recordings were very similar but there was one pronounced difference, which was announced by the

prosecution in Opening Statement and again before the witness Larry Fairfax testified, that there was the sound of a box of tic-tac hard candy shaking in the pocket of the CS close to the recording device.

7. There was a sound on the recording played in court that was not on the recording that I had carefully listened to on March 8, 9 and 10, 2011. That sound was the rattling of tic-tac candy shaking in its box as the CS was moving. No such sound was on the recordings I had heard earlier. That sound dissipated and by the end of the recording, it no longer existed. It appeared to me that the tic-tac sound was added to the recording played in court for the jury after the copies I listened to on March 8, 9 and 10th were made, which tic-tac rattling sound would tend to mask other sounds that the jury might hear, as I heard, and identify, as I had identified, as evidence of non-authenticity.

Further affiant sayeth naught.

I declare under penalty of perjury that the foregoing is correct on this 16th day of July, 2011.



Cyndi G. Steele

The reference to the Tic Tacs is located on the 4/27/2011 transcript on page 303, lines 12-24, which was during the prosecution's opening statement by Mr. Hawes. It reads as follows:

11 And I'm going to tell you right now
12 there is another disturbing piece of evidence for
13 the government, and that is, there is an annoying
14 box of Tic Tacs. You know the little Tic Tacs,
15 the little candies, little plastic box?
16 Mr. Fairfax has a medical condition and
17 his throat gets dry -- kind of like mine is right
18 now. And so he uses Tic Tacs to be able to keep
19 his throat moist. It's a medical condition. So
20 he always has these Tic Tacs. Unfortunately, the
21 Tic Tacs were still on him, and the microphone
22 picks up the Tic Tacs. You're going to hear more
23 about Tic Tacs than you want to hear when you hear
24 those recordings.

Again on page 309, lines 12-14 as follows:

12 on him the same as the day before. And
13 unfortunately, you're going to hear the Tic Tacs
14 again.

Then when LF was being questioned by Whelan, on page 512, lines 4-10

3 BY MS. WHELAN:

4 **Q.** Did you have Tic Tacs in your pocket
5 that day?

6 **A.** Yes, I did.

7 **Q.** Was the recorder in your pocket?

8 **A.** Yes, it was.

9 **Q.** That noise that we hear, is that the
10 Tic Tacs?

11 **A.** Yes, it is.