AMERICAN BOARD of RECORDED EVIDENCE -- VOICE COMPARISON STANDARDS

Abstract

This document specifies the requirements of the American Board of Recorded Evidence for the comparison of recorded voice samples. These standards have been established for all practitioners of the aural/spectrographic method of voice identification and are intended to guide the examiner toward the highest degree of accuracy in the conduct of voice comparisons. These criteria supersede any previous written, oral, or implied standards, and became effective in 1998.

Foreword

This document was developed by members of the American Board of Recorded Evidence, a board of the American College of Forensic Examiners, following their meeting in San Diego, CA in December, 1996. The document draws upon previously published material from the International Association for Identification, the International Association for Voice Identification, The Journal of the Acoustical Society of America, The Audio Engineering Society and The Federal Bureau of Investigation for much of its content. The contents of this document are for non-commercial, educational use. It is the intent of the Board to publish this document in the official journal of the American College of Forensic Examiners.

VOICE COMPARISON STANDARDS

Table of Contents

1. SCOPE

This standard specifies recommended practices for the handling, preparation and analysis of recorded evidence to be followed by practitioners of the aural/spectrographic method of speaker identification. The document covers specific instructions for the preparation of exemplar recordings, voice spectrograms and aural comparison samples. It defines criteria to be applied when arriving at conclusions that are based upon the oral evidence. It also includes requirements for reports and testimony that are offered by the expert witness regarding his findings in voice analyses.

This standard is intended as a guide based upon good laboratory practices for handling recordings that may be used in evidence. Persons handling evidence recordings should first obtain and follow the rules of the legal jurisdiction or jurisdictions involved. When a jurisdiction provides instructions, those should be followed. Only in the absence of such instructions should the recommendations of this standard be followed with the approval of the jurisdiction.

2. EVIDENCE HANDLING.

Since evidence involved in criminal or civil proceedings must meet the appropriate jurisdiction's Rules of Evidence, it is important to properly identify and safeguard it from the time of receipt until returned to the contributor or court. The ABRE has adopted as its standard for handling evidence the AES Standard "AES27-1996 - AES recommended practice for forensic purposes-Managing recorded audio materials intended for examination". The complete document is available at:

Audio Engineering Society, Inc.
60 East 42nd Street
New York, NY 10165

3 PREPARATION OF EXEMPLARS. The quality of the exemplars is critical in allowing an accurate comparison with unknown voice samples.

3.1 Production. The exemplars can be prepared by either the investigator, attorney, examiner, or other appropriate person. Whenever possible, an impartial individual knowledgeable of the known speaker's voice should be present to minimize attempts at disguise, changes in speech rate, adding or deleting accents, and other alterations. The known speaker should state his or her name at the beginning of the recording and repeat the unknown speaker's statement(s) from three (3) to six (6) times, depending upon the length of the unknown samples. Normally, the person preparing the exemplar should record his or her name and that of any other witnesses present.

3.2 Duplication of Recording Conditions.

3.2.1 Microphone. Whenever possible, the same type of microphone system should be utilized when recording exemplars as was used for the original unknown recording. Therefore, if the unknown caller used a telephone, the exemplar should be prepared by having the suspect talk into one telephone instrument and be recorded at a second telephone set, located an appropriate distance away.

3.2.2 Acoustic environment. The exemplar recordings should be prepared in a quiet environment with relatively short reverberation times. Do not imitate noises present at the location of the unknown call or obvious reverberant effects.

3.2.3 Transmission line. Whenever possible, the same general type of transmission line, such as a telephone call, should be utilized when recording exemplars as was used for the original unknown recording.

3.2.4 Recording system. A good quality recording system should always be used in preparing exemplars; it is usually not necessary to imitate the system utilized in recording the unknown sample, but if the system is available and functional, it may be used. A standard cassette set at 1 7/8 inches per second or open reel tape recorder at 3 3/4 or 7 1/2 inches per second or a digital recorder should otherwise be used. Micro cassette and other miniature formats, speeds below 1 7/8 inches per second, and poor quality/inexpensive units are not recommended. Before the known speaker is allowed to leave the exemplar-taking session, the recordings should be played back to insure that the samples are of high quality and properly prepared.

3.2.5 Recording media. Good quality tape or other appropriate recording media should always be used in preparing exemplars; it is not necessary to duplicate the type of tape utilized in recording the unknown sample. The tape should either be new (preferred) or properly bulk erased.

3.3 Duplication of Speech Delivery.

3.3.1 Reading v. recitation. The suspect should be allowed to review the written text or transcription before actually making the recorded exemplars. This familiarity will usually improve the reading of the text and response to oral prompts and increase the likelihood of obtaining a normal speech sample. When a suspect cannot or will not read normally, it is advisable to have someone recite the phrases in the same manner as the unknown speaker and have the suspect repeat them in a similar fashion. Ideally, the exemplar should be spoken in a manner that replicates the unknown speaker, to include speech rate, accent (whether real or feigned), hoarseness, or any abnormal vocal effect. The individual taking the sample should feel free to try both reading and recitation, until a satisfactory exemplar is obtained.

3.3.2 Repetition. Multiple repetitions of the text are necessary to provide information about the suspect's intraspeaker variability. All material to be used for comparison should normally be read or recited from three (3) to six (6) times, unless very lengthy.

3.3.3 Speech rate. Exemplars should be produced at a speech rate similar to the unknown voice sample. In general, the suspect is instructed not to talk at his or her natural speaking rate if this is markedly different from the unknown sample. An effort should be made through repetition to appropriately adjust the speech rate and cadence in the exemplar to that in the questioned recording.

3.3.4 Stress/Accents. Stress includes the emphasis and melody pattern in syllables, words, phrases, and sentences. If prominent or peculiar stress is present in the questioned recording, exemplars should be obtained in a similar manner, if possible. Spoken accents or dialects, both real and feigned, should be emulated by the known speaker. The recitation mode is the better technique for accomplishing this.

3.3.5 Effects of alcohol or other drugs. Since the degree and type of effects from alcohol and other drugs varies from person to person, an attempt to duplicate these vocal changes is not recommended when obtaining the exemplar. If the suspect appears to be under the effects of alcohol or other drugs at the time of the exemplar recording the session should be rescheduled.

3.3.6 Other. If any other unique aural or spectrally displayable speech characteristics are present in the questioned voice, attempts should be made to include them in the exemplars.

3.4 Marking. Same as Sect. 2

4 PREPARATION OF COPIES.

4.1 Playback of Evidential Recordings. The proper playback of the unknown and known voice sample is critical, since it provides the optimum output for the aural and spectral analyses.

4.1.1 Track determination. In situations where the questioned recording was made on equipment of unknown origin or configuration, it may be necessary to analyze oxide on the recording before playing it back. The recorded track position and configuration may be determined by applying an appropriate ferrofluid to the oxide side of analog tapes in a high amplitude portion of the

recording. The treated area is then viewed under low magnification to determine the track configuration and offsets.

4.1.2 Azimuth alignment. Where there is evidence of an audio level or clarity problem during playback, azimuth alignment should be checked and adjusted if necessary by either an inspection of the developed magnetic striations (see track determination above), frequency analysis of the recorded material, or adjustment of the reproducer head azimuth for maximum high frequency output. All audio miniature cassettes, standard cassettes, and open reels (other than loggers) recorded at 15/16 inches per second (2.4 centimeters per second), or less, should be carefully examined for loss of higher frequency information, which often occurs in these formats.

4.1.3 Speed accuracy. Errors in playback speed will cause corresponding variations in the voice frequency, both aurally and spectrally. The playback speed error should be determined for all recordings containing known discrete tones, and then corrected on a reproducer with speed-adjustment circuitry. A Real-Time (RT) Analyzer or Fast Fourier Transform (FFT) analyzer system should be used that allows a resolution of 1% (+0.60 hertz) or better at 60 hertz. Where a known signal is present on the recording, a frequency counter may be employed to correct tape speed. Ideally, there should be less that a 3% error between questioned and known samples that are being compared.

4.1.4 Reproducer. Using the information gleaned from the examinations of the track, azimuth alignment, and speed, a high-quality playback device is configured to allow optimum output.

4.2 Direct Copies. The following information is provided for the analog reel copies that are needed for processing on the Voice Identification, Inc., Series 700 sound spectrograph. If the spectrograph being utilized has a digital memory, the requirements for cabling and retention are still applicable. Even with digital memory systems, a high quality digital or analog tape copy should still be prepared and maintained.

4.2.1 Format. All copies are prepared in a full track, 7 1/2 inches per second format on 1.0 mil or thicker audio tape from a reputable manufacturer. Normally, new, unused reels of tape should be utilized; however, previously recorded tape can be used if either bulk erased or over-recorded on a full track recorder with no input.

4.2.2 Cabling. All copies must be prepared with good quality cables from the playback device to the line input of the recording unit. No loudspeaker-to-microphone copying procedures are permitted.

4.2.3 Recording unit. A separate professional reel recorder, or the one incorporated in the Series 700 Series Spectrograph, is required. At least once a

year, the recorder must be checked by a technically competent individual to determine the unit's playback speed accuracy, distortion level, flutter, record/playback frequency response, and record level. The recorder must meet the following criteria: playback speed within 0.15% distortion of less than 3% at 200 nWb/m, wow and flutter below 0.15% (NAB unweighted), record/playback frequency response of 100 to 10,000 hertz + 3 decibels at 200 nWb/m, and a 0 VU level no greater than 250 nWb/m. If the recorder does not meet all of these standards, it must be repaired and/or adjusted. If a digital system is utilized by the examiner, the system should be checked at least once a year by a technically competent individual according to the manufacturer's written instructions. Digital systems should have almost unmeasurable speed errors, wow and flutter, distortion, and frequency deviations.

4.2.4 Retention. The direct copies must be retained at normal room temperatures and humidity for at least three (3) years, unless the case has been completely adjudicated or the contributor requires the return of all materials used by the examiner.

4.3 Enhanced Copies. When the original recording contains interfering noise and/or limited frequency response, enhanced copies may provide improved audibility and more usable spectrograms. At times, separate enhanced copies will have to be prepared for the aural and spectral examinations to provide optimum results for each. The following information is specifically provided for the analog reel copies that are needed for processing on the Voice Identification, Inc., Series 700 sound spectrograph. If the spectrograph being utilized has a digital memory, the requirements for cabling and retention are still applicable. Even with digital memory systems, a high quality digital or analog tape copy should still be prepared an maintained. A written record of the settings on the devices used should be maintained.

4.3.1 Equalizers. Parametric or graphic equalizers can boost and attenuate selected frequency bands to normalize the recorded speech spectrum. Though an FFT or RT analyzer is of considerable assistance in adjusting the spectrum, a final decision on the equalizer settings should be made by either listening and/or preparing spectrograms, depending upon the enhanced copy's use.

4.3.2 Notch filters. These devices allow the selected attenuation of discrete tones present in the recordings. An FFT or RT analyzer is of considerable assistance in identifying the frequency of the tones and optimally centering the filter's notch.

4.3.3 Deconvolutional filters. These digital devices both automatically attenuate sounds correlated longer than a specified time and flatten the sound spectrum. The filter can, at times, provide improved spectrographic and aural samples for examination. Care should be taken to insure that the adaptation rate is not set at a value that starts to delete speech information.

4.3.4 Other filters. Band pass, shelving, comb, user-characterized digital, and other filters are helpful in a small number of voice identification cases.

4.3.5 Format. Same as 4.2.1.

4.3.6 Cabling. Same as 4.2.2.

4.3.7 Recording unit. Same as Section 4.2.3.

4.3.8 Retention. Same as Section 4.2.4.

5 PRELIMINARY EXAMINATION.

A preliminary examination is conducted to determine whether the unknown and known voice samples meet specific guidelines to allow continuation of the examination.

5.1 Original/Duplicate Recordings. The unknown and known voice samples must be original recordings unless listed as a specific exception below. Copies not meeting these guidelines cannot be used for examination. Short time restraints imposed by the contributor are not considered an exception. When access to the original recording is denied due to legal restraints, copies may be used under the allowed exceptions. The exceptions for not examining the original recordings are:

a. If the original recording has been erased or destroyed, the examiner should then use the best first-generation copy available;

b. The copies were prepared by a qualified voice identification examiner or other technically competent individual following Section 4 guidelines;

c. If the original recording is in a relatively unique format or part of a digital storage system, the examiner or other technically competent individual should prepare the copies from the original material following Section 4 guidelines. If that is not possible, then detailed telephonic and/or written instructions should be given to the individual preparing the copies. Copies produced by non-technical individuals should be closely analyzed in the laboratory to insure that the duplication process was properly done.

5.2 Verbatim/Non-verbatim. The known, or another unknown voice sample, must be either wholly verbatim (preferred), or partially verbatim to allow meaningful comparisons with unknown voice samples. A partially verbatim sample should include phrases and sentences containing at least three (3) similar, consecutive matching words. An example of the use of partial verbatim samples would be two (2) unknown recorded false fire alarms containing, at times, nearly identical

phraseology. If no verbatim recordings are submitted by the contributor, the examiner may analyze the unknown samples to determine whether they would meet the guidelines if appropriate known voice samples are submitted at a later time.

5.3 Number of Comparable words. There must be at least (10) comparable word between two (2) voice samples to reach a minimal decision criteria. Similarly spoken words within each sample can only be counted once. It is noted that in most voice samples at least some of the words identified at this point will not be useful in the final examinations.

5.4 Quality of Voice Samples. This preliminary aural and spectral review is to determine if the voice samples are of sufficient quality to allow meaningful comparisons between them.

5.4.1 Disguise. Samples, or portions of samples, that contain falsetto, true whispering (in contrast to low amplitude speech), or other disguises that obviously change or obscure the vocal formants or other speech characteristics, may need to be eliminated from comparison consideration. Other types of disguise may or may not be usable, depending upon the nature of the disguise. Sometimes a known voice sample with the same type of disguise can be compared, but the examiner should exercise caution in such examinations.

5.4.2 Distortion. Samples, or portions of samples, that include high-level linear and/or nonlinear distortion should be eliminated from comparison consideration. Such distortion can result from saturation of magnetic tape or overdriven electronic circuits, and can produce artifacts, including formants that did not exist in the original speech information.

5.4.3 Frequency range. Samples, or portions of samples, that are restricted in upper frequency range and produce less than two complete speech formants are of limited value to the examiner. Samples producing three or more speech formants provide the examiner better information with which to make a comparison. Sometimes the use of enhanced copies can allow the frequency range to be extended but note the limitations in Section 7.1.3.

5.4.4 Interfering speech and other sounds. Samples, or portions of samples, that contain any extraneous speech information or sounds which interfere with aural identification or spectral clarity should be eliminated from comparison consideration unless the sounds can be sufficiently attenuated through enhancement procedures.

5.4.5 Signal-to-noise ratio. Samples, or portions of samples, containing recording system or environmental noise that impedes aural identification or spectral clarity should be eliminated from comparison consideration unless the noise can be sufficiently attenuated through enhancement procedures.

5.4.6 Variations between samples. Though the following variations can quickly end a voice comparison, the problem can often be remedied by obtaining additional known samples:

a. Transmission systems. Normally, samples being compared should be produced through the same type of transmission system, for example, the telephone, a microphone in a room, or a RF transmitter/receiver. If aurally or spectrally the samples are noticeably different due to the dissimilarities in the transmission systems and filtering does not rectify these differences, no further comparisons should be made.

b. Recording systems. Normally, samples being compared should be produced on either good quality, or compatible, recording systems. However, if the recordings contain uncorrectable system differences that affect aural and spectral characteristics, no further comparisons should be made. Examples of recording differences that can affect the results include high-level flutter, gross speed fluctuations, and voice-activated stop/starts.

c. Speech delivery. Normally, samples being compared should have the speakers talking in the same general manner, including speech rate, accent, similar pronunciation, and so on. However, in cases where this has not been done, as in poorly produced known exemplars, no further comparisons should be made.

d. Other. Any other differences between the voice samples that noticeably effect aural and spectral characteristics should be closely reviewed before proceeding with the examination.

6 PREPARATION OF SPECTROGRAMS.

6.1 Sound Spectrograph. The examiner must use a sound spectrograph, or a digital system, that allows the identification and marking of each speech sound on the spectrogram by either manual manipulation of the drum while listening to the recorded material or the separate identification of the individual sounds on a computer monitor. Spectrographs used must be of professional manufacture, such as the Voice Identification 700 Series or professional computerized systems, such as the Kay Elemetrics Model 5500. The spectrograph should be calibrated at least every six (6) months according to the manufacturer's instructions.

6.1.2 Print Quality. Spectrographic prints must be produced either in an analogue format or, if from a computerized system, must be printed with a minimum of 600 dots per inch resolution.

6.2 Format.

6.2.1 Filter bandwidth. A 250 to 300 hertz bandwidth filter is recommended for the production of most spectrograms. A 450 to 600 hertz bandwidth filter may sometimes improve the formant appearance for high-pitched voices. Narrower filters should only be used for non-voiced sounds and calibration purposes.

6.2.2 Mode. The bar display mode must be used for all spectrograms with the high-shaping equalizer engaged (except when an enhanced copy is being used that has already properly shaped the spectrum).

6.2.3 Frequency range. An appropriate frequency range should be chosen that fully displays all speech sounds in the unknown voice sample. The known voice spectrograms are then prepared using the same frequency range.

6.2.4 Direct v. enhanced. When enhanced copies are used for the examination, at least some spectrograms must be prepared from the direct copies.

6.3 Marking. Each spectrogram must be marked below each speech sound, either phonetically, orthographically, or a combination of both. Great care should be taken to insure that the speech sounds are accurately designated as to how they were spoken, which may not be their correct pronunciation. The spectrograms should be appropriately labeled with identifying information such as specimen, case, and laboratory identifiers. The spectrograms may be marked consecutively for each unknown and known sample. Known and unknown sounds may be marked in different colored ink to facilitate comparisons.

6.4 Retention. All spectrograms should be retained for at least three (3) years after completion of the examination, unless the case has been completely adjudicated or the contributor requires the return of all materials used by the examiner.

7 SPECTROGRAPHIC/ AURAL ANALYSIS.

7.1 Pattern Comparison.

7.1.1 Intraspeaker consistency. The examiner must visually compare similarly spoken words within each voice sample to determine the range of intraspeaker

variability. If there is considerable variability, the word must not be used for comparison. If there is considerable variability in a number of words in a sample, the sample should not be used for comparison. This is often encountered with disguised voices and known exemplars from uncooperative individuals.

7.1.2 Similar speech sounds. Only speech sounds of similarly spoken words should be compared between voice samples. Comparison of the same speech sound but in different words, should be avoided.

7.1.3 Direct v. enhanced. When using spectrograms from direct and enhanced copies, both should be visually compared to words from the known or questioned voice sample. The examiner should be cognizant that the enhancement process may distort the spectral energy distribution, thus increasing the likelihood of a false elimination.

7.1.4 Number of comparable words. This is determined by the total number of different words present in both samples that meet the standards set forth in Section 5.4.1 - 6. A similar or nearly similar word appearing more than once in one or both samples should be counted only as one comparable word.

7.1.5 Speech characteristics.

a. General formant shaping and positioning. A formant is a band of acoustic energy produced by spoken vowels and resonant consonants. Formants and other vocal patterns produced on the spectrograms are visually compared by the examiner. Generally, the spoken word will produce a set or sets of three (3) or more observable formants. A good pattern match exists when the majority, if not all, of the formant shaping and positioning exhibit strong similarities. A precise photographic match rarely occurs even between two (2) consecutive utterances of the same word spoken by the same individual. Conversely even very different voices can exhibit similarities in general formant shaping and positioning for some words. Examination of these patterns must be conducted between each comparable word of the voice samples.

b. Pitch striations. Pitch, or fundamental frequency, can be a useful characteristic for distinguishing between speakers. Pitch information is displayed on a spectrogram in the form of closely-spaced vertical striations, with the spacing and shaping being useful parameters of the individual talker. Differences in the pitch rate and the smoothness or coarseness of the pitch quality should be examined both spectrally and aurally; but most talkers are characterized by fairly wide pitch ranges.

c. Energy distribution. Energy distribution of certain vocal sounds can assist the examiner in analyzing similarities and differences between voice samples. Certain phonemes are displayed primarily by their energy distribution diffused across a certain frequency range. Plosive and fricative consonants are displayed

along the frequency axis as concentrated dark energy distribution patterns. Although the characteristics of energy distributions, especially bursts, are more dependent upon the type of sounds produced than the speakers, some talker-dependent characteristics can be observed.

d. Word length. The time length of a particular spoken word can be readily compared between voice samples. When a person speaks more slowly or faster than normal, the time between words is usually more affected than the length of the individual words. It is noted that a word appearing at the end of a sentence or phrase is usually longer than the same word appearing in the middle.

e. Coupling. The effects of inappropriate coupling can often be observed in spectrograms as either diminished or enhanced energy in the frequency range between the first and second formants. Coupling is related to the open/close condition of the oral and nasal cavities. In normal speaking the nasal cavity is coupled to the oral cavity for nasal sounds, such as "n", "m", and "ng". However, some talkers are hyper nasal, producing nasal-like characteristics in inappropriate vocal sounds; other speakers are hypo nasal producing limited nasal qualities even when appropriate.

f. Other. Plosives, fricatives, and inter-formant features should be spectrally compared between samples by the examiner. Other sounds such as inhalation noise, repetitious throat clearing, or utterances like "um" and "uh" can sometimes be compared to the known exemplar if they have been successfully replicated.

7.2 Aural Comparison.

7.2.1 Short-term memory. An aural short-term memory comparison must be conducted either by playing the two (2) samples on separate playback systems with a patching arrangement to allow rapid switching between them or by recording short phrases or sentences from each sample on the same recording. The short-term memory playback tape should contain all words used in the spectrographic comparison. The two (2) samples should be reviewed at approximately the same speech amplitude and with the same general frequency range. The frequency range may be normalized between the samples by using band pass filtering on the sample with the widest frequency range to duplicate the range found on the other sample.

7.2.2 Direct v. enhanced. When direct and enhanced copies have been produced, both should be aurally compared to the known or questioned sample. The examiner should recognize that though enhancement procedures often improve intelligibility, they can also produce changes, at times, that can make samples of the same talker sound somewhat different.

7.2.3 Pronunciation. Only similarly pronounced words should be compared between samples.

7.2.4 Intraspeaker consistency. The examiner must aurally compare similar words within each sample to determine if they are spoken in a generally consistent manner. If intraspeaker variability is present for a particular word, that word should not be compared to the other voice sample. If considerable intraspeaker variability is present in the entire sample, that sample should not be used for comparison. This is often the problem with disguised speech and known exemplars from uncooperative individuals.

7.2.5 Speech characteristics.

a. Pitch. See sect. 7.1.5.b.

b. Intonation. Intonation is the perception of the variation of pitch, commonly known as a melody pattern. Spontaneous conversation will normally exhibit this characteristic to a greater extent than a passage that is read by the speaker.

c. Stress/Emphasis. The stress or emphasis within the words of the sample should be similar for different recordings of the same talker when no disguise is present.

d. Rate. The rate of speaking under the same conditions is relatively constant for a particular talker. However, rates of reading, recitation, and conversation will normally vary for the same talker.

e. Disguise. Obvious vocal disguises can disqualify a sample for comparison purposes. The examiner should carefully analyze the characteristics of the disguise in a sample and then determine if it is possible to make a meaningful comparison with another sample, whether it also contains a disguised voice or not.

f. Mode. Certain speaker-dependent characteristics can be discerned from the mode in which a speaker initiates sounds. Speakers range from gradually to abruptly initiating voicing, which can reveal useful similarities and differences between two samples.

g. Psychological state. Listening usually reveals many of the effects of an altered psychological state upon the voice. Alterations may be characterized as nervousness, over-excitement, excessive monotone, crying, and so on. The examiner should be cautious in comparing samples with major changes due to an altered psychological state.

h. Speech defects. Speech defects are abnormalities in the voicing of sounds, and can include lisps, pitch and loudness problems, and poor temporal sequencing. Except for extreme cases, there are no criteria to assess whether a voice is considered normal or defective. Obvious, or even subtle, defects in the

questioned or known voice samples can often provide vital information in the comparison decision.

i. Vocal quality. Vocal quality is the perception of the complex, dynamic interplay of the laryngeal voicing (pitch, intonation, and stress), articulator movement, and oral cavity resonances. Since each individual's voice is relatively unique in its vocal quality, comparisons can provide important information regarding similarities and differences between the voice samples.

j. Other. Examples of other useful speech characteristics that are occasionally heard include long-term fluctuations of pitch (vibrato), vocal fry (extremely low pitching), pitch breaks, and stuttering.

7.3 Conclusions. Every aural/spectrographic examination conducted can only produce one of seven (7) decisions; Identification, Probable Identification, Possible Identification, Inconclusive, Possible Elimination, Probable Elimination, or Elimination. The following descriptions for each decision are the minimal decision criteria, and must be adhered to by the examiner, except that lower confidence level can always be chosen, even though the criteria would allow a higher degree of confidence. Within the range of probable decisions, the examiner may wish to clarify his findings, i.e. low probability, high probability, depending upon the quantity and quality of the comparable material available to the examiner. Comparable words must meet the previously listed criteria. The following are the seven (7) possible decisions.

7.3.1 Identification. At least 90% of all the comparable words must be very similar aurally and spectrally, producing not less than twenty (20) matching words. Each word must have three (3) or more usable formants. This confidence level is not allowed when there is obvious voice or electronic disguise in either sample, or the samples are more than six (6) years apart.

7.3.2 Probable Identification. At least 80% of the comparable words must be very similar aurally and spectrally, producing not less than fifteen (15) matching words. Each word must have two (2) or more usable formants.

7.3.3 Possible Identification. At least 80% of the comparable words must be very similar aurally and spectrally, producing not less than (10) matching words. Each word must have two (2) or more usable formants.

7.3.4 Inconclusive. Falls below either the Possible Identification or Possible Elimination confidence levels and/or the examiner does not believe a meaningful decision is obtainable due to various limiting factors. Comparisons that reveal aural similarities and spectral differences, or vice versa, must produce an Inconclusive decision.

7.3.5 Possible Elimination. At least 80% of the comparable words must be very dissimilar aurally and spectrally, producing not less than (10) that do not match. Each word must have two (2) or more usable formants.

7.3.6 Probable Elimination. At least 80% of the comparable words must be very dissimilar aurally and spectrally, producing not less than fifteen (15) words that do not match. Each word must have two (2) or more usable formants.

7.3.7 Elimination. At least 90% of all the comparable words must be very dissimilar aurally and spectrally, producing not less than twenty (20) words that do not match. Each word must have three (3) or more usable formants. This confidence level is not allowed when there is obvious voice or electronic disguise in either sample, or the samples are more than six (6) years apart.

7.4 Second Opinion. A second opinion is not required, but may be obtained from another certified examiner when desired by either the examiner or the party submitting the evidence.

7.4.1 Independence. A second opinion must be completely independent of the first examiner's decision, and no oral or written information shall be provided regarding that first opinion.

7.4.2 Material provided. The second examiner should only be provided the originals, or direct and enhanced copies, any work notes under Sections 2, 3, and 4 and the spectrograms. The second examiner must not be provided any materials that reflect even partially, the first examiner's opinions regarding the examination.

7.4.3 Examination. A thorough analysis should be conducted by the second certified examiner, using the guidelines in Sections 5, 6 and 7 (except for 7.4). It is left to the discretion of the second examiner whether to prepare additional spectrograms or copies.

7.4.4 Resolving differences. If different decisions are reached by the two (2) examiners, a detailed discussion between them of the analysis will often lead to a resolution. If not, the lower confidence level must be reported and testified to when both decisions are an identification or elimination. If split between and identification and elimination, no matter what the confidence level, the decision must be inconclusive. A third independent decision can be obtained but the result will be the lowest confidence level, or an inconclusive of all the examiners involved.

7.4.5 Reporting. Whenever possible, the second examiner should prepare a short report listing the results of the second opinion. This is not necessary if both examiners are in the same organization. The name and results of the second opinion can then be included in the first examiner's work notes.

8 WORK NOTES.

8.1 Required Information. The examiner's work notes should be in accordance with Rule 26 of the Federal Rules of Evidence - Expert Witness Statement categories, and should contain, as a minimum, the following information:

a. Laboratory, case, and specimen identifiers;

b. Description of submitted evidence;

c. Chain-of-custody documentation;

d. Track determination, azimuth alignment, and speed accuracy information, where required, for each submitted sample;

e. Information on the duplication processes, including the type of equipment and format copies;

f. Information of the enhancement processes, if any, including the type of equipment, filter settings, and format copies;

g. List of the exact words used for comparison and whether they matched or not;

h. Name of any second opinion examiner and the results of that examination;

i. Final decision.

8.2 Retention. The work notes should be retained for at least three (3) years after completion of the examination unless the contributor has requested that all material relating to the case be returned.

9 REPORTING.

9.1 Format. The report should be typed, dated, and in a standard laboratory or business letter style. The content of the report should be in conformity with Rule 26 of the Federal Rules of Evidence. The following information must be included: a short description of the evidence being examined, a summary of the examination performed, the final decision, and a statement of accuracy. Exhibits, handouts and supporting documentation should be separate from the report. Business matters, such as payment of fees, should be set forth in separate communications and not included within the report.

9.2 Decision Statement. The report must clearly state which of the seven (7) decision options listed in Section 7.3 was the final result of the examination.

10 TESTIMONY.

The American Board of Recorded Evidence does not take a position as to whether or not a certified examiner should provide testimony regarding examination results. However, an examiner must follow the standards set forth in this document, including the appropriate criteria set forth in this section, whether they provide testimony, or not.

10.1 Testimony v. Investigative Guidance. Each specific organization or individual examiner must decide before conducting spectrographic voice identification examinations whether testimony will be provided. If not, the contributor must be advised of the investigative guidance policy and all oral and written reports should set forth this information.

10.2 Qualification List. The presentation of the qualifications of the examiner should be in conformity with Rule 26 of the Federal Rules of Evidence - Expert Witness Statement categories, regarding expert witnesses.

10.3 Pre testimony Conference. Discussion of the examination with the attorney before judicial proceedings is an important aspect of providing meaningful testimony and educating the attorney on the strengths and limitations of the technique. The conference should include a candid discussion, the inherent problems, identification of scientific literature that is either critical or supportive, and other information important to the testimony.

10.4 Appearance and Demeanor. Whenever possible, examiners must dress in proper business attire or appropriate law enforcement or military uniform for all judicial proceedings, maintain a professional demeanor even under adversarial conditions, and direct explanations to the jury, when present.

10.5 Presentation. The examiner should provide to the judge and/or jury, as a minimum, his/her qualifications, an overview of the spectrographic technique, its scientific basis, the details of the analysis procedures followed in the specific case, and the results of the analysis. The information should be presented in a form understandable to non-experts, but with no loss of accuracy.