Introduction

This OSAC Speaker Recognition process map arose from the need to establish a common frame of reference to help overcome differences in terminology and participants' background and experience. The development of the process map helped the participants to better understand current practices and communicate them in a constructive way.

Representatives of multiple U.S. government agencies, individual practitioners, and international experts met for three days with a facilitator to create the first draft that sketched the components of a forensic examination. The current version incorporates additional contributions from a variety of researchers and practitioners. The OSAC Speaker Recognition subcommittee would like to acknowledge and thank all those who participated in the development of this process map.

The process depicted does not represent the practice of any single laboratory, but generalizes the diverse practices of multiple laboratories. This document reflects a balance between an attempt to be comprehensive and the efficient use of volunteers' time. It is intended to be descriptive only, and its release does not imply endorsement by the OSAC Speaker Recognition Subcommittee of any specific approach or process. No inferences should be drawn from the inclusion or exclusion of any approach or process or from the level of detail provided for any particular approach or process.

This process map is not intended to represent a best practice but rather to facilitate the development of future best practice documents by the OSAC Speaker Recognition Subcommittee.
1000 - Administrative Assessment

1100 Case Suitability

1200 Case Acceptance

2000 - Technical Assessment

2100 Preliminary Evaluation

2200 Content Review

3000 - Processing

3100 Pre-Analysis Observations and Processing

3200 Relevant Population Data

4000 - Analysis

4100 Select analysis method(s)

4200 Human Supervised Automatic Method

4300 Holistic Auditory Perceptual

4400 Expert-Driven Auditory Phonetic & Acoustic Phonetic Analysis

4500 Spectrographic

4600 Blind Grouping

5000 - Results

5100 Evaluation/Generating Conclusion

5200 Verification

5300 Case Close-Out

Terminate Case

Legend

Process start/end
Multistep subprocess
Selection from multiple options
1100 - Case Suitability

1110 Case Acceptance Criteria
- Approval received?
- Analysis timeframe appropriate?
- Technically feasible?
- Sufficient quality?
- Sufficient quantity?
- Knowns collected properly?
- Analysis requested appropriate?
- Resources available to address bias?

1130 All case acceptance criteria met?
- Y To 1200
- N

1200 - Case Acceptance

1210 Request Received?
- Y To 1215
- N

1220 Assign Case Priority
- Y To 1225
- N

1230 Log Chain of Custody
- Y To 1235
- N

1240 Will foils (imposters) be used in this examination?
- Y To 1250
- N

1250 Are there appropriate foils available in this case?
- Y To 1260
- N

1260 Use foils from case
- Y To 1270
- N

1270 Assign Case
- Y To 1280
- N

1280 Will case be worked/assessed by more than one examiner?
- Y To 1290
- N

1290 Transfer Case to Assigned Examiner ensuring that no information about the case has been given to secondary analyst.

Commentary

Legend
- Single process step
- Multistep process that may be pre-defined in a standard, by lab policy, and/or by examiners
- Decision step
- Indicates that the next or previous step is somewhere else on the process map

This document is a work product of the OSAC Speaker Recognition Subcommittee.
2100 - Preliminary Evaluation

2110 Determine Physical Controls and Data Handling Requirements (Gloves, virus checks, write protection, etc) Take steps to protect the evidence (make a working copy)

2120 Inventory (mark and describe evidence per agency policy)

2130 Optimizing Playback (finding the best way to playback media) (see SWGDE Best Practices for Forensic Audio)

2140 Best Available Data?  
Y N

2150 Communicate with Submitter or check archives to obtain better samples  
Y N

2160 Can I proceed with current or obtained samples?  
Y N

2200 - Content Review

2210 Best Available Data?  
Y N

2220 Communicate with Submitter or check archives to obtain better samples  
Y N

2230 Can I proceed with current or obtained samples?  
Y N

2240 Expert Consult Needed? (translator, video analyst, etc)  
Y N

2250 Consult Expert  
Y N

2260 Can I continue with examination?  
Y N

Legend

- Single process step
- Multistep process that may be pre-defined in a standard, by lab policy, and/or by examiners
- Decision step
- Indicates that the next or previous step is somewhere else on the process map

This document is a work product of the OSAC Speaker Recognition Subcommittee.
3100 - Pre-Analysis Observations and Processing

3105 Select question of interest for analysis
3110 Document intrinsic and extrinsic properties in Q and K (Q is the limiting evidence)
3115 Document intrinsic and extrinsic properties in Q and K
3120 Do speaking style and recording conditions match between Q and K?
3125 Document the matching and mismatching conditions and features of the Q and K, as applicable
3130 Do I have a method that works with these data?
3135 Will processing permit a method selection?
3140 Process Evidence
3145 Select appropriate method(s)
3150 Will processing improve performance of method(s)?
3155 Process Data (see list of processing techniques)
3160 Did processing elucidate any new intrinsic or extrinsic properties?
3165 Document newly observed properties
3170 Do I have appropriate data to carry out downstream processing?
3175 Can I obtain or simulate comparable test data?
3180 Obtain or simulate required data
3190 Continue to method (consider contextual bias when deciding which method to use first)
4100 – Select Analysis Method(s)

4110 Assess suitability of Human Supervised Automatic Analysis

4120 Assess suitability of Holistic Auditory Perceptual Analysis

4130 Assess suitability of Expert-Driven Auditory Phonetic and Acoustic Phonetic Analysis

4140 Assess suitability of Spectrographic Analysis

4150 Assess suitability of Blind Grouping Method

4115 Is this method appropriate?

To 4200

4125 Is this method appropriate?

To 4300

4135 Is this method appropriate?

To 4400

4145 Is this method appropriate?

To 4500

4155 Is this method appropriate?

To 4600

Commentary

Acoustic Phonetic Statistical Analysis (Semiautomatic Analysis)

Acoustic Phonetic Statistical Analysis (or Semiautomatic Analysis) is similar to Human Supervised Automatic Analysis (4200), but uses features derived via phonetic analysis, including human-supervised measurements of acoustic properties of the speech recording.
4200 - Human Supervised Automatic Analysis

Process Map of Current Practices in Forensic Speaker Recognition

4205 Is diarization needed?
4210 Diarize speakers in recording
4215 Optimize the selection
4220 Create new file containing one speaker per recording
4225 Have I separated all speakers of interest?
4227 Can I continue with examination? (data sufficiency)
4230 Conduct expert critical listening (set expectations and parameter selection)
4235 Do I have a system optimized for my case data?
4240 Can I optimize my automated speaker recognition system for the case conditions?
4241 Train system using training data
4242 Test system using development data
4243 Performance adequate?
4245 Perform validation test?
4246 Test system using validation data
4247 Report validation results
4248 Performance adequate?
4249 Apply system to known-speaker and questioned-speaker recordings
4250 Post-process results?
4255 Report only numeric result (e.g. score, likelihood ratio)?
4256 Adjust model/reference population and/or settings.
4260 Is the output logically consistent with regards to type of mismatch?
4265 Stop analysis?
4267 Document issues noted in box 4260
4270 Adjust model/reference population and/or settings.
4275 Report that system performance under the case conditions is inadequate to proceed with an evaluation of the strength of evidence associated with the known- and questioned-speaker recordings
4280 Report results
4285 Convert result to verbal description
4287 Assess confidence in ASR results
4290 Complete Human Supervised Automatic Method? (formerly known as another ASR?)

From 4100 To 5000

Commentary

This document is a work product of the OSAC Speaker Recognition Subcommittee.
4500 – Spectrographic / Aural Spectrographic Analysis

4510
Conduct examination according to standard (e.g. IAI Standard, reference JFI (1991) 41:5)

4520
Document results

Legend

<table>
<thead>
<tr>
<th>Icon</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single process step</td>
<td></td>
</tr>
<tr>
<td>Multistep process</td>
<td>Multistep process that may be pre-defined in a standard, by lab policy, and/or by examiners</td>
</tr>
<tr>
<td></td>
<td>Indicates that the next or previous step is somewhere else on the process map</td>
</tr>
</tbody>
</table>
Process Map of Current Practices in Forensic Speaker Recognition

Legend

- Single process step
- Multistep process that may be pre-defined in a standard, by lab policy, and/or by examiners
- Process step that results in documentation
- Indicates that the next or previous step is somewhere else on the process map

4600 – Blind Grouping Method

Researcher #1

4605: Are there appropriate foils available in the case?

- Yes
  - 4610: Create a list of speakers to be included in the blind grouping, containing at least each Q, and K and G or more fails.
  - 4615: Is diarization needed?
    - Yes
      - 4620: Diarize speakers in recordings
    - No
      - 4625: Optimize the selection
        - Create new file per speaker per recording. Strip the recording from any speech that contains name, addresses, etc. that would help linking fragments.
      - 4630: Is it possible to make 1 or more fragments of max 20 seconds for each speaker and each recording?
        - Yes
          - 4635: Is it possible to make true same speaker and true different speaker pairs?
            - Yes
              - 4640: Create the max 20 fragments into separate audio files, all with the same audio specs (take the one with the highest quality).
            - No
              - 4645: Create a list of fragments per speaker per recording, max 20 fragments.
        - No
          - 4650: Anonymize and collect all the fragments in one larger audio file in random order (but make a look-up table to retrieve the origins). Present the larger file to researcher #2.
          - 4655: Document a grouping, including perceived similarity within groups and perceived dissimilarity between groups and the ungrouped fragments.

- No
  - 4660: Anonymize and collect all the fragments in one larger audio file in random order (but make a look-up table to retrieve the origins). Present the larger file to researcher #2.

Researcher #2

4670: Is each of the fragments suitable for grouping?

- Yes
  - 4675: Compare each of the fragments with each other and group them by perceived speaker identity.
  - 4680: Document a grouping, including perceived similarity within groups and perceived dissimilarity between groups and the ungrouped fragments.
- No
  - 4685: Separate out the unsuitable fragments.
  - 4690: Interpret grouping using look-up table and including correctness of grouping in ground truth.
  - 4695: Document results

4690: Interpret grouping using look-up table and including correctness of grouping in ground truth.

Researchers #1 and #2

4690: Interpret grouping using look-up table and including correctness of grouping in ground truth.

4695: Document results

Commentary

To Next Step
**5100 – Evaluation / Generating Conclusion**

- **5105** (Single process step): Multiple methods used?
  - **Y**: 5115 Assess confidence for each method used
  - **N**: 5110 Assess confidence

- **5115** (Decision step): Categorize (least) result
  - **Y**: 5120 Assign weight and combine (least) results per agency policy
  - **N**: 5125 Formulate conclusion

- **5125** (Single process step): Final case (least) result?
  - **Y**: 5150 Will consultation be helpful?
  - **N**: 5155 Share case detail Seek input

- **5150** (Decision step): Will additional analysis be helpful?
  - **Y**: 5160 Document Opinion or Conclusion
  - **N**: 5170 Are all case questions answered?

- **5160** (Single process step): Draft preliminary report

**Legend**

- Single process step
- Multistep process that may be pre-defined in a standard, by lab policy, and/or by examiners
- Decision step
- Process step that results in documentation
  - Indicates that the next or previous step is somewhere else on the process map

---

**5200 - Verification**

- **5205** (Single process step): Verify conclusions (according to agency policy)
  - **Y**: 5210 Conduct blind review
  - **N**: 5220 Conduct technical review (per agency policy)

- **5210** (Decision step): Did the report need to be changed?
  - **Y**: 5225 Share limited case data (per agency policy)
  - **N**: 5230 Review conflicts (Follow agency conflict resolution policy)

- **5225** (Single process step): Does the report need to be changed?
  - **Y**: 5240 Conduct independent analysis (examiner begins at 525)
  - **N**: 5250 Review technical (per agency policy)

- **5230** (Decision step): Will consultation be helpful?
  - **Y**: 5245 Seek input
  - **N**: 5255 Share case detail

- **5240** (Single process step): Share case detail 

---

**5300 - Case Close-Out**

- **5305** (Single process step): Notify requestor and transmit report, where appropriate

---

This document is a work product of the OSAC Speaker Recognition Subcommittee.
### Process Step

#### 1000 – Administrative Assessment

#### Description

#### Terms and Definitions

#### Comments

#### Issues

#### References

---

This document is a work product of the OSAC Speaker Recognition Subcommittee.
## 1100 – Case Suitability

### Description

The description of the case suitability process step is not provided.

### Terms and Definitions

The terms and definitions for case suitability are not provided.

### Comments

The comments on case suitability are not provided.

### Issues

The issues related to case suitability are not provided.

### References

The references for case suitability are not provided.

---

**Revised**

September 30, 2019 11:22 AM
<table>
<thead>
<tr>
<th>Process Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1200 – Case Acceptance</td>
<td></td>
</tr>
</tbody>
</table>

Terms and Definitions

Comments

Issues

References

Revised
September 30, 2019 11:22 AM
<table>
<thead>
<tr>
<th>Process Step</th>
<th>2000 – Technical Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>Terms and Definitions</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>Issues</td>
<td></td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
</tbody>
</table>

Revised

September 30, 2019 11:22 AM
<table>
<thead>
<tr>
<th>Process Step</th>
<th>2100 – Preliminary Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>Terms and Definitions</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>Issues</td>
<td></td>
</tr>
<tr>
<td>References</td>
<td>SWGDE Best Practices for Forensic Audio.</td>
</tr>
</tbody>
</table>

Revised
September 30, 2019 11:22 AM
<table>
<thead>
<tr>
<th>Process Step</th>
<th>2200 – Content Review</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Terms and Definitions</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Comments</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Issues</strong></td>
<td></td>
</tr>
<tr>
<td><strong>References</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Revised</strong></td>
<td>September 30, 2019 11:22 AM</td>
</tr>
</tbody>
</table>
### Process Step

#### 3000 – Processing

<table>
<thead>
<tr>
<th>Process Step</th>
<th>Description</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Terms and Definitions</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Comments</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Issues</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Revised</th>
</tr>
</thead>
</table>

September 30, 2019 11:22 AM

This document is a work product of the OSAC Speaker Recognition Subcommittee.
### Term and Definitions

3155: Types of processing

- Enhancement for intelligibility or listenability (e.g. tone removal, spectral shaping, adaptive filtering, etc.)
- Normalization
- Convert sampling rate / bit depth
- Channel conversion
- DC offset
- Anti-aliasing

### Comments

### Issues

### References

### Revised

September 30, 2019 11:22 AM
<table>
<thead>
<tr>
<th>Process Step</th>
<th>3200 – Relevant Population Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>This block describes the necessary use of different data sets for testing system performance. Evaluation typically requires training data and test data sets, but other data may be required (e.g. for calibration).</td>
</tr>
</tbody>
</table>

**Term ands Definitions**

**Comments**

**Issues**

**References**


**Revised**

September 30, 2019 11:22 AM
<table>
<thead>
<tr>
<th>Process Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000 – Analysis</td>
<td></td>
</tr>
</tbody>
</table>

**Terms and Definitions**

**Comments**

**Issues**

**References**


**Revised**

September 30, 2019 11:22 AM
### 4100 – Select Analysis Method(s)

<table>
<thead>
<tr>
<th>Description</th>
<th>Terms and Definitions</th>
<th>Comments</th>
<th>Issues</th>
<th>References</th>
</tr>
</thead>
</table>

### Revised

September 30, 2019 11:22 AM
Process Step
4200 – Human Supervised Automatic Analysis

Description
This block describes the necessary use of different data sets for testing system performance. Evaluation typically requires training data and test data sets, but other data may be required (e.g. for calibration).

Terms and Definitions
4230: “Expert critical listening” is defined as ...

Comments

Issues

References

Revised
September 30, 2019 11:22 AM
<table>
<thead>
<tr>
<th>Process Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>4300 – Holistic Auditory Perceptual Analysis</td>
</tr>
</tbody>
</table>

### Description

- This section details the process of holistic auditory perceptual analysis in forensic speaker recognition.

### Terms and Definitions

- Here, terms related to auditory perceptual analysis are defined.

### Comments

- Space for comments or notes related to the process step.

### Issues

- Potential issues or concerns related to the process step.

### References

- Relevant references or literature supporting the process step.

---

This document is a work product of the OSAC Speaker Recognition Subcommittee.
## Process Step

### 4400 – Expert Driven Auditory Phonetic and Acoustic Phonetic Analysis

### Description

**Terms and Definitions**

**Issues**

### References


### Revised

September 30, 2019 11:22 AM

This document is a work product of the OSAC Speaker Recognition Subcommittee.
4500 – Spectrographic / Aural Spectrographic Analysis

Description

Terms and Definitions

Comments

Issues

References


Kersta L.G. (1962). Voiceprint identification. *Nature*, 196, pp. 1253–1257. [http://dx.doi.org/10.1038/1961253a0](http://dx.doi.org/10.1038/1961253a0)


<table>
<thead>
<tr>
<th>Process Step</th>
<th>4600 – Blind Grouping Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td></td>
</tr>
<tr>
<td>Terms and Definitions</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>Issues</td>
<td></td>
</tr>
<tr>
<td>Process Step</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
</tbody>
</table>

**Terms and Definitions**

**Comments**

**Issues**

**References**

Revised

September 30, 2019 11:22 AM
Process Step

S100 – Evaluation / Generating Conclusion

Description

Terms and Definitions

Comments

References


Revised

September 30, 2019 11:22 AM

This document is a work product of the OSAC Speaker Recognition Subcommittee.
<table>
<thead>
<tr>
<th>Process Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5200 – Verification</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Terms and Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Revised</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 30, 2019 11:22 AM</td>
</tr>
</tbody>
</table>

This document is a work product of the OSAC Speaker Recognition Subcommittee.
### Process Step

<table>
<thead>
<tr>
<th>Process Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5300 – Case Close-Out</td>
<td></td>
</tr>
</tbody>
</table>

### Terms and Definitions

<table>
<thead>
<tr>
<th>Terms and Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

### Comments

<table>
<thead>
<tr>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

### Issues

<table>
<thead>
<tr>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

### References

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

---

This document is a work product of the OSAC Speaker Recognition Subcommittee.