Special Session on Forensic Acoustics
162nd Meeting of the Acoustical Society of America
San Diego, 3 November 2011
http://sandiego2011.forensic-acoustics.net/

Defining the default defense hypothesis in likelihood-ratio forensic voice comparison

Felipe Ochoa
Geoffrey Stewart Morrison

FORENSIC VOICE COMPARISON LABORATORY
SCHOOL OF ELECTRICAL ENGINEERING & TELECOMMUNICATIONS

UNSW
THE UNIVERSITY OF NEW SOUTH WALES
SYDNEY • AUSTRALIA

Acknowledgment of Funding

- Research supported by the Australian Research Council, Australian Federal Police, New South Wales Police, Queensland Police, National Institute of Forensic Science, Australasian Speech Science and Technology Association, and the Guardia Civil through Linkage Project LP100200142. Unless otherwise explicitly attributed, the opinions expressed are those of the authors and do not necessarily represent the policies or opinions of any of the above mentioned organizations.

$$LR = \frac{p(\text{Evidence} \mid H_{\text{prosecution}})}{p(\text{Evidence} \mid H_{\text{defense}})}$$

$H_{\text{prosecution}}$ = same speaker

H_{defense} = different speaker

- A forensic likelihood ratio as a strength-of-evidence statement is the answer to a specific question.

- The question is in part defined by the defense hypothesis.

- The defense hypothesis is usually more specific than "some other speaker on the planet".

- In DNA-profile comparison, the samples are submitted blind.

- In forensic voice comparison the submitter has listened to the offender recording and decided that it sounds sufficiently similar to the suspect that it is worth submitting for forensic analysis

  (i.e., forensic comparison of the suspect and offender recording).

- They have generated the same-speaker hypothesis.

- The default defense hypothesis should therefore be:

  The suspect is not the same speaker as on the offender recording but is one member of a population of speakers whom to the submitter sound sufficiently similar to the offender recording that the submitter would submit recordings of these speakers for forensic analysis.

- A background sample is used to model the distribution of acoustic properties in the relevant population.

- A scientifically defensible procedure for selecting speakers to include in the background sample could be to have a panel of listeners select recordings which sound sufficiently similar to the offender recording that they would submit them for forensic analysis.

- The panel of listeners should listen under conditions as similar as possible to those under which the same-speaker hypothesis was originally generated.

- Categories such as gender and accent are not relevant.

- The panel of listeners should be as similar as possible to the listener who originally generated the same-speaker hypothesis.

Three casework examples:

Channel Mismatch

- Western Australia v Mansell, WA Dist Ct, No 665 of 2008
  - Morrison (2010, §99.1040–99.110)

- a police officer listened to a series of telephone-intercept recordings
  - over a period of several days
  - mobile telephone

- later part of a team conducting a search of a suspect's office
  - heard the voice of someone out of sight talking with one of her colleagues
  - immediately recognized it as the same as the voice on the intercepts

Channel Mismatch

- recordings of the suspect were made at the time of the search and later
- prosecution did not submit suspect and offender recordings for forensic analysis
  – instead relied on police officer’s testimony
- defense called expert to testify on validity of non-technical speaker identification
- suspect was acquitted

Channel Mismatch

- How should a background database have been selected?

- no dispute that the offender was a male Australian-English speaker
  - larger data set from which panel of listeners selected could be male Australian-English speakers

- listen to offender recordings over several days
  - mobile telephone

- listen to and select potential background recordings
  - audio quality as close as possible to original listening conditions
    - potentially high-quality audio
    - background noise and room acoustics

Channel Mismatch

- select recordings which given the channel mismatch sound sufficiently similar to offender recordings that they would submit them for forensic analysis

Male or female?

- 2009 case for the Forensic Acoustics and Audiovisual Section of the Central Forensic Science Laboratory of the Chilean Investigative Police
- suspect was an adult male with, for an adult male, a high pitched voice
- offender recording had a superficially similar pitch
- Should the background database have consisted of
  - males?
  - males with high pitched voices?
  - females?
  - some combination of the above?

Male or female?

- The questions about gender were misplaced.

- **Error:** Conditioning background database on suspect
  - we do not know the identity of the offender
    - we do not know the gender of the offender
  - the defense hypothesis is that the suspect is not the offender
    - knowing the gender of the suspect does not tell us anything about the gender of the offender

- **Error:** Focus on gender as a category
  - The panel of listeners should select recordings of speakers which sound sufficiently similar to the offender recording that they would submit them for forensic analysis.

Accent

- California v Prinzivalli 1985
  - Labov & Harris (1994)

- Pan Am executives in Los Angeles received threatening phone calls
  - identified the voice as that of a Pan Am employee

- Labov was asked to analyze recordings of the suspect and offender
  - suspect had a New York City accent
  - offender had a Boston accent

- Prosecutor insisted on continuing
  - Ladefoged and Disner presented spectrographic evidence

- Judge heard prosecution summation
  - acquitted

Accent

- From a likelihood-ratio perspective:
  - Some speakers can speak with more than one accent; however, (pending any evidence that this is the case for the suspect) the probability of getting the two different accents on different recordings is much higher if they were produced by different speakers than if they were produced by the same speaker.

- What kind of speakers should be included in the background sample?
  - speakers with New York City accents?
  - speakers with Boston accents?

Accent

- The panel of listeners should be from Southern California and not be familiar with North East accents, and should select recordings of speakers which to them sound sufficiently similar to the offender recording that they would submit them for forensic analysis.
  - some may have New York accents
  - some may have Boston accents
  - some may have other accents

- accent as a category is not relevant

- The panel of listeners should listen under conditions as similar as possible to those under which the same-speaker hypothesis was originally generated.

- Categories such as gender and accent are not relevant.

- The panel of listeners should be as similar as possible to the listener who originally generated the same-speaker hypothesis.

Thank You