

A demonstration of the evaluation of forensic evidence under conditions reflecting those of an actual forensic-voice-comparison case

Ewald Enzinger^{1,2} and Geoffrey Stewart Morrison^{1,3}

¹*School of Elec. Eng. & Telecom., University of New South Wales, Sydney, Australia*

²*Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria*

³*Department of Linguistics, University of Alberta, Edmonton, Canada*

ewald.enzinger@oeaw.ac.at, geoff.morrison@forensic-evaluation.net

This presentation demonstrates the evaluation of forensic evidence under conditions reflecting those of an actual forensic-voice-comparison case. This includes consideration of the relevant prosecution and defence hypotheses to address in this case, selection of data reflecting the adopted defence hypothesis, simulation of recording conditions reflecting those of the suspect and offender recordings in the case, quantitative measurement and statistical modelling to calculate a likelihood ratio given the relevant hypotheses and under recording conditions reflecting those of the case, and empirical testing of the validity and reliability of the resulting system given the relevant hypotheses and under recording conditions reflecting those of the case. As such, this provides a practical demonstration of a forensic voice comparison conducted under a paradigm which we have previously espoused (see Morrison, 2013, and Morrison & Stoel, 2013, for recent summaries of the paradigm).

There was no dispute in this case that the suspect and the speaker on the offender recording were adult male Australian English speakers, and we were able to draw samples from a database of multiple non-contemporaneous recordings of adult male Australian English speakers. The database included high-quality recordings of speech from an information-exchange-via-telephone task and a face-to-face interview task, which best reflected the speaking styles in the offender and suspect recordings respectively. We will discuss how the defence hypothesis in this case was further refined from adult-male Australian English speaker and how a relevant subset of the database was selected.

The offender recording in this case was of a landline telephone call made to a call centre. As well as telephone transmission, it included background noise at the call centre, and it was saved in a compressed format. The suspect recording was of a police interview conducted in a reverberant room with ventilation noise and saved in a compressed format. The presentation will include a description of how we simulated these conditions so that database recordings could be converted and used to train and test statistical models under conditions reflecting those of the case. We will play audio recordings which illustrate the steps in simulating the recording conditions.

The presentation will also include brief descriptions of: the procedures used to make quantitative measurements of the acoustic properties of the voices on the recordings, statistical modelling procedures used to calculate likelihood ratios (details of channel compensation techniques is the subject of another proposed presentation), and the procedure used to empirically test the validity and the reliability of the system. Finally, the results of system testing will be presented.

References

- Morrison, G. S., & Stoel, R. D. (2013 online). Forensic strength of evidence statements should preferably be likelihood ratios calculated using relevant data, quantitative measurements, and statistical models – a response to Lennard (2013) Fingerprint identification: How far have we come? *Australian Journal of Forensic Sciences*. doi:10.1080/00450618.2013.833648
- Morrison, G. S. (2013 online). Distinguishing between forensic science and forensic pseudoscience: Testing of validity and reliability, and approaches to forensic voice comparison. *Science & Justice*. doi:10.1016/j.scijus.2013.07.004