What really matters in forensic comparison? Validation, decision making, and uncertainty

Vincent Hughes
vincent.hughes@york.ac.uk
Department of Language and Linguistic Science, University of York

Abstract

In 2005, Saks and Koehler described a *paradigm shift* in forensic science, towards the use of more scientifically defensible, systematic, and objective methods for evaluating expert evidence. This has led to the widespread adoption of the likelihood ratio (LR) framework as a measure of the strength of forensic comparison evidence, as well as growing regulatory pressure for experts to empirically validate methods under conditions reflecting those of casework. Such developments have undoubtedly changed the face of many forensic disciplines, not least forensic speech science. Yet, it is important to acknowledge that any forensic comparison, whether empirically validated and data-driven or entirely qualitative and experience-driven, involves subjective decision making on the part of the expert. Such decisions introduce uncertainty, which can affect the reliability of the analysis and the eventual conclusions that an expert presents to an end-user.

In this talk, I will outline a process for method validation, arrived at via a consensus between a group of forensic speech scientists, forensic statisticians, and lawyers (Morrison et al. 2021), which recognises the subjectivity and uncertainty in forensic comparison. I will then focus on two specific and crucial issues: (i) the definition of the relevant population, representing the defence proposition, and (ii) data sampling from that relevant population for building and validating forensic approaches. This will lead me to ask the fundamental question: *what matters most in forensic comparison?*. I will argue that, ultimately, we should be driven by minimising uncertainty, which in turn helps end-users make better decisions, rather than maximising discriminability (despite this generally being the focus of our research). At the very least, this involves experts recognising the decisions and assumptions they make throughout an analysis, and being explicit about those decisions and assumptions such that they are open to scrutiny. I will also discuss ways in which it may be possible to reduce uncertainty, via empirical processes such as calibration.

References

Morrison, G. S., Enzinger, E., Hughes, V., Jessen, M., Meuwly, D., Neumann, C., Planting, S., Thompson, W., van der Vloed, D., Ypma, R. and Zhang, C. (2021) Consensus on validation of forensic voice comparison. *Science and Justice* 61(3): 299-309.

Saks, M. J. and Koehler, J. J. (2005) The coming paradigm shift in forensic identification science. *Science* 309: 892-895.

Biography:

Dr Vincent Hughes is a Lecturer in Forensic Speech Science at the University of York, UK. His research interests lie in forensic speech science, automatic speaker recognition, and forensic data science, as well as phonetics, phonology, and sociolinguistics. His current research focuses on understanding the bases and limitations of speaker-specificity in speech and the extent to which linguistics and automatic speaker recognition can be integrated, particularly for forensic purposes. He was recently awarded an AHRC Early Career Grant to examine the relative performance of human listeners and automatic speaker recognition systems, comparing the conditions under which the different methods outperform one another. Dr Hughes is also interested in the application of the numerical likelihood ratio to the evaluation of forensic evidence and appropriate methods for validating forensic systems. He was part of a group that published a recent consensus on validation of forensic voice comparison. He has also been part of a group writing a best practice manual for forensic voice comparison using linguistic and phonetic methods on behalf of the European Network of Forensic Science Institutes (ENFSI).