

Forensic speaker recognition between linguistics and speech technology

Michael Jessen, Bundeskriminalamt Wiesbaden

Forensic Phonetics and Acoustics (aka Forensic Speech Science or Forensic Speech and Audio Analysis) is concerned with tasks involving spoken language and/or audio signals, embedded in contexts ranging from police investigations to cases presented in court. Several of these tasks involve speaker recognition in a broad sense, subdivided into speaker profiling, forensic voice comparison and voice line-up (not addressed in this presentation). In speaker profiling, a recording of an unknown offender is available, and the main task of the forensic expert is to infer from the recording as many group-level characteristics as possible, including regional variety, social features, foreign accent and an estimation of age. In forensic voice comparison, not only is a recording of the offender available but there is a suspect, and a recording of that suspect can often be obtained. The task of the expert is to provide evidence supporting identity or non-identity of the individuals behind the recordings.

Speaker profiling is a highly linguistic task and it benefits from linguistic knowledge, especially on language varieties and on the level of phonetics and phonology (Schilling & Marsters 2015). Forensic voice comparison is heavy on phonetics, including acoustic phonetics and the analysis of extragrammatical aspects such as global pitch, tempo or voice quality (Nolan 1983; Foulkes & French 2012; Jessen 2012). Since about twenty years, forensic voice comparison increasingly benefits from a method that is derived from speech technology: automatic speaker recognition (Hansen & Hasan 2015). An interesting methodological development in forensic voice comparison that concerns forensic science in general, is the use of Bayesian statistics and the notion of the likelihood ratio (Rose 2002; Jessen 2018; Morrison & Enzinger 2018).

The presentation begins with the illustration of a recent forensic case that started as a speaker profiling task and – after a suspect was found aided by the results of the profile – continued as a forensic voice comparison. The second part of the presentation goes through a number of linguistic and phonetic speaker characteristics that are commonly used, most particularly in forensic voice comparison. The third part gives a very brief presentation of the idea behind Bayesian statistics and illustrates the live operation of an automatic speaker recognition system.

References:

- Foulkes, P. & P. French (2012): Forensic speaker comparison: A linguistic-acoustic perspective. In: L.M. Solan & P.M. Tiersma (eds.), *Oxford handbook of language and law*, 557–572. Oxford: Oxford University Press.
- Hansen, J.H.L. & H. Taufiq (2015): Speaker recognition by machines and humans – A tutorial review. *IEEE Signal Processing Magazine*, Nov. 2015, 74–99.
- Jessen, M. (2012): *Phonetische und linguistische Prinzipien des forensischen Stimmenvergleichs*. München: LINCOM.
- Jessen, M. (2018): Forensic voice comparison. In: J. Visconti (ed. in collab. with M. Rathert), *Handbook of communication in the legal sphere*, 219–255. Berlin: Mouton de Gruyter.
- Nolan, F. (1983): *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press.
- Morrison G.S., E. Enzinger & C. Zhang (2018). Forensic speech science. In: I. Freckelton & H. Selby (eds.), *Expert evidence* (Ch. 99). Sydney: Thomson Reuters.
- Rose, P. (2002): *Forensic speaker identification*. London: Taylor and Francis.
- Schilling, N. & A. Marsters (2015): Unmasking identity: Speaker profiling for forensic linguistic purposes. *Annual Review of Applied Linguistics* 35: 195-214.