

## **The role of regional variation in the quantification of strength-of-evidence for forensic voice comparison**

In forensic voice comparison (FVC) the expert is commonly asked to offer a conclusion on whether the voices in a pair of criminal and suspect recordings belongs to the same or different individuals. The likelihood ratio (LR) has become established as the “logically and legally correct” framework for the evaluation of such comparison evidence (Rose and Morrison 2009: 143) across forensic disciplines (Robertson and Vignaux 1995, Evett 1991). Applied to FVC, the LR involves an assessment of the similarity of features across the samples and their typicality in the *relevant population* (Aitken and Taroni 2004), since strength-of-evidence is dependent on “whether the values found matching (...) are vanishingly rare, or sporadic” (Nolan 2001: 16).

However, the definition of the *relevant population* for FVC casework has largely been overlooked as an issue which may compromise the reliability of numerical LRs. Despite the mass of LVC research identifying that the distribution of phonetic and linguistic features in the population is affected by a multitude social and stylistic factors, generally only sex and language, defined broadly according to country, are controlled (Rose 2004, Morrison 2008, Kinoshita et al 2009) as *logically relevant* (Kaye 2004, 2008) factors for FVC.

This paper presents the results of two studies which assess the extent to which LRs are affected by mismatch between evidential recordings and reference data with regard to regional variety. Using polynomial estimations of formant trajectories from spontaneous GOOSE /u:/ (F1/F2) and PRICE /aɪ/ (F1/F2/F3) tokens, same-speaker and different-speaker LR comparisons were performed on four sets of eight speaker test data where only one set matches the reference data (89-102 speakers) for regional dialect.

GOOSE same-speaker LRs were on average 3.3 times higher for mismatch sets: equivalent to the difference between ‘limited’ and ‘moderate’ support for the prosecution. Mismatch different-speaker comparisons also displayed high levels (58-71%) of contrary-to-fact support for the prosecution. Consistent with a priori expectations about the level of regional variation across the two lexical sets, the divergence between PRICE same-speaker LRs was considerably higher (up to 5.4 times greater for mismatching sets). However, the removal of regionally defining information (F1 and/or F2) was shown to reduce the effect of dialect mismatch. Therefore the extent to which regional variety affects LRs is predictably dependent on the features under analysis. Further, for certain features it may be more *logically relevant* to control for factors such as age, ethnicity and class.

### **References:**

- Evett, I. W. (1991) Interpretation: a personal odyssey. In Aitken, C. G. G. and Stoney, D. A. (eds.) *The use of statistics in forensic science*. London: Ellis Horwood. 9-22.
- Kaye, D. H. (2004) Logical relevance: problems with the reference population and DNA mixtures in *People v. Pizarro*. *Law, Probability and Risk* 3: 211-220.

Kaye, D. H. (2008) DNA probabilities in *People v. Prince*: When are racial and ethnic statistics relevant? In Speed, T. And Nolan, D. (eds.) *Probability and Statistics: Essays in Honour of David A Freedman*. Beachwood, OH: Institute of Mathematical Statistics. 289-301.

Kinoshita, Y., Ishihara, S. and Rose, P. (2009) Exploring the discriminatory potential of F0 distribution parameters in traditional speaker recognition. *International Journal of Speech, Language and the Law* 16(1): 91-111.

Morrison, G. S. (2008) Forensic voice comparison using likelihood ratios based on polynomial curves fitted to the formant trajectories of Australian English /aɪ/. *Journal of Speech, Language and the Law* 15(2): 249-266.

Nolan, F. (2001) Speaker identification evidence: its forms, limitations, and roles. *Proceedings of the 'Law and Language: Prospect and Retrospect' Conference*. 12-15 December 2001, Levi, Finland.

Robertson, B. and Vignaux, G. A. (1995) *Interpreting evidence: evaluating forensic science in the courtroom*. Chichester: John Wiley.

Rose, P. (2004) Technical Forensic Speaker Identification from a Bayesian Linguist's Perspective. Keynote paper, *Forensic Speaker Recognition Workshop, Speaker Odyssey '04*. 31 May - 3 June 2004, Toledo, Spain. 3-10.

Rose, P. and Morrison, G. S. (2009) A response to the UK Position Statement on forensic speaker comparison. *International Journal of Speech, Language and the Law* 16(1): 139-163.