

# PERCEPTUAL TESTING AND SPEECH SYNTHESIS IN THE ANALYSIS OF DISPUTED UTTERANCES

Allen Hirson, David Howard, Elizabeth McClelland  
A.Hirson@city.ac.uk; dh@ohm.york.ac.uk; earsemc@hotmail.com

## ABSTRACT

In cases of forensic speaker identification work, spectrographic and specifically formant analysis are widely used to supplement traditional auditory phonetic analysis. There is, however, less agreement as to the most appropriate methods in relation to the analysis of disputed utterances – or more generally in the transcription of evidential recordings.

In relation to a recent case involving disputed utterances from a mobile telephone clip, the conflicting opinions of experts were backed up on each side by formant measurements. The present study considers two approaches to the resolution of conflicts of opinion over the interpretation of a disputed utterance. (1) Perceptual testing to supplement standard formant analysis. (2) Measured formants synthesized to generate stimulus utterances which are then perceptually tested by phonetically trained listeners.

The relevant section of the mobile telephone video clip was presented in open field to 100 phonetically trained listeners. The listeners were provided (in random order) with the alternative interpretations to the utterance. Alternative interpretations (i.e. corresponding to *neither* of the Experts' views) were also invited. The results of this simple test, including an analysis of the alternative interpretations, are presented.

To test the significance of the (discrepant) formant measurements of key vowel sounds of the disputed utterance, it is also proposed that the latter may be synthesized from the measured formants [1], and the results appropriately grafted into the carrier phrase to generate experimental utterances. The latter may then be evaluated by phonetically trained listeners, and these results compared with the original transcriptions. More generally, results are discussed in terms of the contribution of vowel formant measurements in speech decoding [2]. While neither method explored here may be practicable in all cases of this kind, increasing the repertoire of methods available to the forensic phonetician may be helpful where the application of traditional phonetic methods and formant analysis produces equivocal results.

## References

- [1] Mullen J., Howard D.M. & Murphy D.T. (2007) Real-time dynamic articulations in the 2D waveguide mesh vocal tract model. *IEEE Transactions on Speech and Audio Processing*, 15 (2), 577-585.
- [2] Rosner B.S. & Pickering J.B. (1994) *Vowel Perception and Production*. Oxford University Press.