

WikiDialects: a resource for assessing typicality in forensic voice comparison

Vincent Hughes¹ and Jessica Wormald^{1,2}

¹*Department of Language and Linguistic Science, University of York, UK.*

²*J P French Associates, York, UK.*

vincent.hughes@york.ac.uk

jessica.wormald@jpfrench.com

In forensic voice comparison (FVC), the strength of the evidence is dependent not only on the similarity between the suspect and offender with regard to the features present in the voices, but also the typicality of these features within the wider, relevant population (Aitken and Taroni 2004). Despite calls across forensic science (and from the UK Government Forensic Regulator) for more robust and replicable estimations of typicality, experts still commonly rely on their experience, eminence and intuition when evaluating strength of evidence in FVC casework (for issues with this see Ross et al. 2016). One crucial reason for this is the relative lack of baseline descriptions of regional and, in particular, social varieties of languages. Comprehensive descriptions of varieties and dialectological work are no longer fashionable in sociolinguistics. Rather, papers tend to focus on a single linguistic feature for the purposes of exploring an aspect of variation and/or change. Moreover, these papers are often small in scale and dispersed across a range of journals, books, and blogs. Of the purely descriptive works that are available many are out of date (e.g. Wells 1982) and therefore may provide misleading accounts of varieties, which in turn may affect the conclusions that the expert arrives at in contemporary FVC cases. Finally, despite the large literature in sociolinguistics, there remains insufficient coverage of the diversity of language varieties even within individual countries (e.g. British English).

We are currently working on a funding application to address these issues. Our proposed solution is to create a wiki for descriptions of language varieties. A wiki is an online encyclopaedia (i.e. a knowledge pool) which is used, developed and updated by a community. In this case, the wiki would be a central resource for summaries of, and signposts to, academic works and anecdotal accounts of patterns of language variation within different regional and social groups. The wiki would be a ‘living thing’ in that it would continually be updated as new research is conducted. We envisage that there would be a relatively large community of potential contributors and users. Contributions would primarily come from academics working in sociolinguistics, phonetics, and dialectology. Given the original motivation for this resource, it would be of central value to experts and academics working in forensic speech science. But beyond this, baseline descriptions of language varieties would also be of use to those working in speech and language therapy and language education (e.g. TESOL), and of general interest to dialect societies and the wider public.

We have arranged our grant proposal into three sections:

- (1) **Creating the wiki.** This will involve deciding on the appropriate format of information and ways of indexing according to both regional/social factors and individual linguistic-phonetic features.
- (2) **Populating the wiki.** As an initial step, we intend to collate research from sociolinguistics and dialectology to add to the wiki. This will focus on British English, but the resource will support users adding descriptions of varieties from other countries and languages.
- (3) **Outreach.** We intend to encourage contributions to the wiki through workshops and guest lectures to academics in sociolinguistics, phonetics, and forensics. We will also

publicise the resource to a wider audience, including the general public, in order to maximise its utility.

References

- Aitken, C. G. G. and Taroni, F. (2004) *Statistics and the Evaluation of Evidence for Forensic Scientists (2nd edition)*. Chichester: Wiley.
- Ross, S., French, J. P. and Foulkes, P. (2016) UK practitioners' estimates of the distribution of speech variants. Paper presented at the *International Association of Forensic Phonetics and Acoustics (IAFPA)* conference, University of York, UK. 24th – 26th July 2016.