How defining 'accents' affects forensic speaker comparison

To appropriately evaluate the strength of forensic speaker comparison evidence, it is essential that the expert considers not only the phonetic similarity between the offender and suspect voices samples, but also the typicality of the speech patterns relative to what's called the 'relevant population'. This is usually done by defining group-level properties of the offender: such as age, gender and regional background/influences. These decisions are, in themselves, evidential, since they involve narrowing down the pool of potential offenders in the case. In XXXX (XXXX), we proposed a theoretical framework based firstly on assessing the strength of group-level evidence (i.e. how usual is the accent? Level 1) and then assessing the strength of the individual-level evidence (i.e. how unusual are the speech patterns relative to other speakers of the same accent? Level 2). In this paper we explore how these decisions are made, how we might assess the typicality of an accent and ultimately how all of this affects the strength of evidence.

There are issues with the idea of a person having a singular 'accent', but the pool of potential offenders has to be narrowed down to some extent. There is also some dispute as to whether the court or jurors can decide and assess the accent and its impact on conclusions or whether this falls within the expert's role (XXXX, XXXX). Further, the methods by which experts arrive at descriptions of 'accent' has not been scrutinised in any depth. Ross, French and Foulkes (2016) show that expert's typicality judgments for certain features vary quite drastically, and that different experts rely on differing sources of information (experience, literature, recorded databases, etc).

This paper will describe the results of testing using both simulations and real data which show that poorly defined 'accent profiles' ot typicality judgements for phonetic features can have a considerable effect on the resulting strength of speaker comparison evidence. In the worst cases, poor typicality judgements can produce errors, which, in the real world, could lead to miscarriages of justice. However, there are potential solutions. Projects such as SPADE (Stuart-Smith, et al. 2017) and the 'Dialect Apps' (Leeman, et al. 2016), for example, provide potentially invaluable sources of reference material to assist an expert in arriving at group-level descriptions and typicality assessments. Additionally, qualitative resources can help to signpost experts and other researchers to a wide range of resources and provide overviews of findings, and crucially, highlight where gaps in the research are. Ongoing developments in automatic tools (e.g. Y-ACCDIST, Brown & Wormald 2017) and machine learning also provide exciting opportunities for the future. All of these developments will complement experience-led judgements, and should reduce reliance on outdated reference material; this will increase the validity and objectivity accent descriptions.

Brown, G., & Wormald, J. (2017). Automatic Sociophonetics: Exploring corpora using a forensic accent recognition system. *Journal of the Acoustical Society of America*, 142(1), 422-433.

Gold, E., Ross, S., & Earnshaw, K. (2018). The 'West Yorkshire Regional English Database': Investigations into the generalizability of reference populations for forensic speaker comparison casework. In *Interspeech 2018: Speech Research for Emerging Markets in Multilingual Societies* (pp. 2748-2752).

Leemann, A., Kolly, M. J., & Britain, D. (2016). English Dialect App.

Ross, S., French, J. P., & Foulkes, P. (2016) UK Practitioners' Estimates of the Distribution of Speech Variants. *Presentation; 25th IAFPA conference, York, UK.*

Stuart-Smith, J., Sonderegger, M., & Mielke, J. (2017-2020). SPeech Across Dialects of English (SPADE): Large-scale digital analysis of a spoken language across space and time. ESRC Grant ES/R003963/1, NSERC/CRSNG Grant RGPDD 501771-16, SSHRC/CRSH Grant 869-2016-0006, NSF Grant SMA-1730479. (Digging into Data/Trans-Atlantic Platform).