Diphthong dynamics in unscripted speech

Vincent Hughes[1]*, Kirsty McDougall*[2]*, and Paul Foulkes*[1,3]

[1]*Department of Language and Linguistic Science, University of York, UK*
vh503@york.ac.uk

[2]*Department of Linguistics, University of Cambridge, UK*
kem37@cam.ac.uk

[3]*J P French Associates, York, UK*
pf11@york.ac.uk

A growing body of research (e.g. McDougall 2004, 2006) show that dynamic aspects of formant frequencies carry considerable speaker-distinguishing information and that measuring formant contours leads to greater speaker discrimination than 'static' measurements. Work by McDougall (2004, 2006) argues that the diphthong /aɪ/ is a good candidate for formant dynamic analysis since a diphthong offers large scope for individual variation. In that study, tokens of Australian English /aɪ/ in read speech yielded high levels of speaker-discrimination when formant dynamic measurements were subjected to discriminant analysis.

The present study extends this work to investigate the effectiveness of formant dynamic techniques at distinguishing speakers in unscripted speech. Previous studies of formant dynamics have generally been undertaken using read speech in controlled situations. A small experiment on Standard Southern British English (SSBE) /uː/ (McDougall 2007) showed that although discrimination rates were lower than in read speech, /uː/ formant dynamics in spontaneous speech did yield speaker-distinguishing information. The present study expands on this work by examining /aɪ/ in unscripted speech from 20 speakers of SSBE from the DyViS database (Nolan et al. in press). An average of 14 tokens of /aɪ/ per speaker were extracted from mock police interviews. Segmentation was performed by hand and measurements of the first three formants across the entire diphthong trajectory were extracted via a Praat script. The dataset will be compared using discriminant analysis to assess the extent to which the diphthongs differentiate between speakers, and to compare the performance of the spontaneous data relative to the scripted data reported by McDougall (2006). The effect of consonant transitions on the data will also be examined. Implications for forensic speaker comparison will be discussed.

References

McDougall, K. (2004) Speaker-specific formant dynamics: an experiment on Australian English /aɪ/. *International Journal of Speech, Language and the Law* **11**, 103-130.

McDougall, K. (2006) Dynamic features of speech and the characterisation of speakers: towards a new approach using formant frequencies. *International Journal of Speech, Language and the Law* **13**, 89-126.

McDougall, K. (2007) Distinguishing speakers using formant dynamics in read and

spontaneous speech: a study of British English /uː/. Paper presented at the International Association for Forensic Phonetics and Acoustics Annual Conference, Plymouth, 22-25 July 2007.

Nolan, F., McDougall, K., de Jong, G. and Hudson, T. (in press) The DyViS Database: Style-Controlled Recordings of 100 Homogeneous Speakers for Forensic Phonetic Research. *International Journal of Speech, Language and the Law* **16**.