The correlation structure of speech parameters in Southern Standard British English

Erica Gold[1] *and Vincent Hughes*[1]

[1]*Department of Language and Linguistic Science, University of York*.
{erica.gold|vh503}@york.ac.uk

The research presented in this paper is funded by a research grant from the International Association of Forensic Phonetics and Acoustics. This work builds upon a previous pilot study (Gold and Hughes 2012) and explores the correlation structure of spontaneous speech from a sociolinguistically-homogeneous set of speakers using a series of segmental, suprasegmental and linguistic parameters.

Data were extracted from a subset of speakers from the Dynamic Variability in Speech (DyViS) database (Nolan et al., 2009) and consist of:

- midpoint F1, F2 & F3 values for FLEECE (/iː/), TRAP (/a/), & NORTH (/ɔː/)
- midpoint F1, F2 & F3 values hesitation markers UM and UH
- dynamic F1, F2 & F3 values for PRICE (/aɪ/)
- long-term formant distributions (LTFD) F1-F4
- mean and standard deviation of fundamental frequency (f0)
- mean articulation rate (AR)
- voice onset time (VOT) for word-initial /t/ and /k/
- click rate (the number of velaric ingressive stops per minute)

Mean values were calculated for each speaker for each element of each variable, and a correlation matrix generated based on pairwise Spearman correlation coefficients. Pairwise correlation tests were conducted for each individual speaker and patterns compared with those of the group. Finally, Euclidean distances between variables were generated based on speaker means for each element using multidimensional scaling, as a means of developing a graphical model for all of the linguistic-phonetic variables analysed.

In terms of group patterns, a number of theoretically predictable correlations were found. There is a high degree of dependence between mean F3 values across all of the tested vocalic parameters. Similarly, a negative correlation was found between mean VOT of /t/ and mean AR. Both of these correlations are predicted by linguistic theory. The between-speaker correlation tests also revealed non-significant relationships between parameters which were expected to be correlated (f0 and F1), as well as unexpected significant correlations, such as that between mean click rate and LTFD2 (p=0.028). Interestingly, when considering patterns of correlation across elements of the same phoneme such as that between F2 and F3 of UM, a strong positive correlation was found for group means, despite some sets of within-speaker values displaying no correlation or even a negative correlation between F2 and F3.

The results highlight the overall complexity of the correlation structure of linguistic-phonetic variables as well as the extent to which this complexity is predicted by phonetic theory and the degree of agreement across within- and between-speaker correlations. The implications for combining analyses of individual speech variables into an overall assessment of the strength of evidence will be explored for both LR-

and non LR-based forensic speaker comparison.

References

Gold, E. and Hughes, V. (2012) Defining interdependencies between speech parameters. Poster presented at *the Bayesian Biometrics for Forensics (BBfor2) Summer School in Forensic Evaluation and Validation*, Universidad Autonoma de Madrid, Madrid. 18th - 21st June 2012.

Nolan, Francis, McDougall, Kirsty, de Jong, Gea and Hudson, Toby. (2009) The DyViS database: style-controlled recordings of 100 homogeneous speakers for forensic phonetic research. *International Journal of Speech, Language and the Law*, 16(1): 31–57.