Articulation rate as a discriminant in forensic speaker comparisons

Erica Gold[1*]

[1]Department of Language and Linguistic Science, The University of York, York, United Kingdom

*erica.gold@york.ac.uk

Forensic phoneticians have found speech tempo to be an important parameter for forensic speaker comparisons, with 93% of experts analyzing speech tempo (Gold and French 2011). It is also reported that 20% of all experts identified speech tempo as one of the most useful parameters for discriminating speakers, ranking it as the third most helpful parameter (alongside F0) out of all possible parameters used in a forensic speaker comparison. With regards to forensic phonetics, speech tempo is typically quantified as either speaking rate (SR) or articulation rate (AR). The majority of forensic phoenticians prefer AR over SR (Gold and French 2011), because AR has been found to have lower intra-speaker variability than SR (Künzel 1997).

This study presents the analysis of AR and standard deviation (SD) for 100 Southern Standard British English male speakers. The results examine both the inter- and intra-speaker variability of AR in spontaneous speech, as well as assess the evidential value of AR as a parameter in forensic speaker comparisons.

AR data were collected from 26-32 local AR tokens per speaker, with 2993 local AR measurements taken in total. Results showed that there was a higher degree of intra-speaker variation than inter-speaker variation, and roughly 68% of the population falls within a 1.3 syllables/second range.

To assess the evidential value of articulation rate, likelihood ratios (LRs) were computed using a MatLab implementation of Aitken and Lucy's (2004) Multivariate Kernel-Density formula (Morrison 2007). It was found that AR performs much better with same speaker (SS) comparisons than different speaker (DS) comparisons. For DS comparisons, the system is performing slightly worse than chance (46% compared to 90% correct in SS comparisons). Additionally, severity of error was assessed using log-LR cost (Cllr). The Cllr for AR is .8981 which is close to 1, but still under; classifying it as a 'poor' performance (van Leeuwen and Brümmer 2007).

Following this study it appears that AR may not be the discriminant shibboleth all experts hope for. However, it is important that AR is still considered in forensic speaker comparisons in conjunction with other speech parameters. As Rose (2006) points out, "not all speakers differ from each other in the same". Therefore, there will be individuals where AR is a highly discriminant parameter for them, as is evident for those individuals in the 100 DyVis speakers who fell in the outer ranges of the population distribution.

References

Aitken, C.G.G. and Lucy (2004) Evaluation of trace evidence in the form of multivariate data. *Applied Statistics 54*: 109-122.

Gold, E. and French, P. (2011) International practices in phonetic speaker comparison. *International Journal of Speech, Language and the Law* 18.2: 293-307.

Künzel, H. (1997) Some general phonetic and forensic aspects of speaking tempo. *Forensic Linguistics* 4:48-83.

Morrison, G.S. (2007) MatLab implementation of Aitken and Lucy's (2004) forensic likelihood ratio software using multivariate-kernel-density estimation. Downloaded: December 2011.

Nolan, F. (2009) *Dynamic Variability in Speech: a Forensic Phonetic Study of British English, 2006-2007* [Studio Task 2]. Colchester, Essex: UK Data Archive [distributor], July 2011. SN: 6790.

Rose, P. (2006). The intrinsic forensic discriminatory power of diphthongs. *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*. 6-8 December 2006, University of Auckland, New Zealand. 64-69.

van Leeuwen, D.A. and Brümmer, N. (2007). An introduction to application-independent evaluation of speaker recognition systems. In Müller, C. (ed.) *Speaker Classification I*, LNAI 4343. Berlin: Spinger-Verlag, 330-353.