Examining articulatory settings using MRI: pilot results

XXX
xxx

xxx

Oral presentation preferred

Introduction

The Vocal Profile Analysis (VPA) scheme (Laver, 1991) is well established as a tool for analysing voice quality, and is used for forensic speaker identification and speech therapy applications. The VPA scheme comprises perceptual ratings of articulatory settings, muscular tension, and phonation features, which are rated by experts on the basis of the audio signal alone. As the ratings are based on the perception of the rater, there can be discrepancies between experts' ratings of the same speech, and use of the scheme requires a training and calibration process (San Segundo *et al.*, 2018).

In this project, we directly measure the articulators using magnetic resonance imaging (MRI) during speech produced with different articulatory settings, with reference to the VPA scheme. These data permit a greater understanding of how the perceptual ratings of the VPA scheme correlate with real articulator positions, and quantify the differences between different articulatory settings, allowing articulatory strategies to be compared across speakers. Furthermore, the results can be used to inform phonetics teaching, and to produce materials for the training and calibration of VPA ratings.

Method

Expert speakers are placed in an MRI scanner and asked to produce a range of utterances in the following 'extreme' articulatory settings, in addition to a neutral setting: close/open jaw, pharyngeal constriction, raised/lowered larynx, backed/fronted tongue body, lip rounding/protrusion and lip spreading. Dynamic, two-dimensional data for the midsagittal plane are collected for running speech, from which landmarks are extracted to provide physical measurements associated with each setting. Additionally, three-dimensional images of the whole vocal tract are collected for several held vowels, permitting the measurement of cavity volumes and the construction of a detailed three-dimensional model of the subject's vocal tract, capable of reproducing intermediate VPA degrees in addition to extreme settings.

Audio recordings are captured simultaneously with the MRI data, using an optical microphone. Due to the noise in the MRI scanner, separate audio recordings are also captured in 'MRI-like' conditions (Gully *et al.*, 2019) inside an anechoic chamber, to provide an additional source of high-quality audio for monitoring and comparison. These recordings are rated by two experts using the VPA scheme, in order to ensure that the target articulatory settings are produced throughout the study.

Findings

Results will be presented for a single subject, including quantitative data on the relative articulator positions, cavity volumes, and associated measurements. Where appropriate, these will be compared with data for other subjects, particularly where articulatory strategies vary. Results will form part of a larger study, the data from which will be made available for research once processing and analysis is complete.

References

Gully, A. J., Foulkes, P., French, J. P., Harrison, P. T., and Hughes, V. (2019). The Lombard effect in MRI noise. *Proceedings of the International Congress of Phonetic Sciences 2019* [in press].

Laver, J. (1991). *The Gift of Speech: Papers in the Analysis of Speech and Voice*. Edinburgh: Edinburgh University Press.

San Segundo, E., Foulkes, P., French, P., Harrison, P., Hughes, V., and Kavanagh, C. (2018). The use of the Vocal Profile Analysis for speaker characterization: methodological proposals. *Journal of the International Phonetic Association*, 1-18. doi:10.1017/S0025100318000130.