Investigating articulatory setting using MRI: preliminary results

Amelia Gully[1], Paul Foulkes[1], Peter French[1,2], Philip Harrison[1,2], Vincent Hughes[1]

[1]*Department of Language and Linguistic Science, University of York, York, U.K.*
[2]*J. P. French Associates, York, U.K.*

Articulatory setting refers to how a speaker habitually holds their supralaryngeal vocal tract articulators, giving a long-term voice 'quality' across speech events, independent of utterance. This is commonly quantified on an auditory basis, using schemes such as the vocal profile analysis (VPA) protocol [1]. However, to date very little work has been done to match auditory measures of articulatory setting with measured differences in articulation. This study seeks to answer the following questions in relation to this goal: 1) how well does the physical articulation of each setting match the auditory description?; 2) to what degree are other articulators involved in producing each setting?; and 3) do different subjects produce settings in a similar way, once differences in vocal tract size and shape are accounted for?

In order to address these questions, two professional phoneticians (one male, one female) were asked to read a short passage and hold several vowels (/i, a, ɑ, ɔ, u, ə/) in a magnetic resonance imaging (MRI) scanner. Utterances were spoken in a neutral setting, and then with each of eight different settings that would result in an 'extreme' (but non-pathological) VPA score: fronted tongue body, backed tongue body, raised larynx, lowered larynx, open jaw, closed jaw, lip rounding and lip spreading. Audio data were captured during the scans using an optical microphone, and separately using high quality recording equipment in an anechoic chamber after the MRI session (the anechoic recording environment was designed to be as close as possible to MRI-like conditions, see [2]). The data reported here are based on midsagittal MRI slices for the held vowel /ə/ only.

Using the techniques of geometric morphometrics [3], between-subject differences in vocal tract size and shape were normalised, allowing a detailed comparison of articulatory settings. This technique permits visualisation of the vocal tract shape under each setting as a deformation grid from the mean vocal tract shape. The results reveal that for most settings, both subjects used a similar articulatory strategy to achieve the target, but these were more similar in some settings than others and in the case of lip rounding, differed substantially. Additionally, certain settings produced vocal tract shapes that were further apart in the shape space, indicating different articulatory strategies; whereas other settings were more closely clustered, indicating similar articulatory strategies for different acoustic targets.

References

[1] J. Laver. *The Gift of Speech: Papers in the Analysis of Speech and Voice*. Edinburgh: Edinburgh University Press, 1991.
[2] A. J. Gully, P. Foulkes, J. P. French, P. T. Harrison, and V. Hughes, "The Lombard effect in MRI Noise," in *Proc. ICPhS 2019*, Melbourne, Australia, 2019, pp. 800-804.
[3] D. C. Adams, M. L. Collyer, and A. Kaliontzopoulou, *Geomorph: Software for geometric morphometric analyses*. R package version 3.1.0. Available: https://cran.r-project.org/package=geomorph, 2019.