Assessing midsaggital tongue contours in polar coordinates using generalised additive (mixed) models

Ultrasound tongue imaging (UTI) is a non-invasive technique that enables researchers to study, more or less directly, the articulation of the tongue during speech at appreciably high resolution. On the other hand, the quantitative analysis of UTI data remains more of a challenge. Since the publication of the seminal paper by Davidson (2006), statistical modelling of whole tongue contours has been mostly dominated by the use of Smoothing Splines Analysis of Variance (SSANOVA). Mielke (2015) and Heyne & Derrick (2015) further showed that fitting SSANOVA with tongue data in polar coordinates (rather than cartesian coordinates) produces more reliable results. SSANOVA of tongue contours undoubtedly advanced our understanding of tongue articulation and speech modelling. However, SSANOVA comes with some limitations, e.g. separate models are fitted for different phonetic contexts even within a single speaker, and the model implementation does not include random effects. On the other hand, developments in the statistical world have seen an increased use of Generalised Additive Models (GAMs) and their mixed-effects counterpart (GAMMs, Wood, 2006). This family of models is a highly flexible solution which extends standard generalised linear mixed models for the modelling of non-linear effects.

This paper will offer a review of GAMMs fitted to tongue contours in polar coordinates. An introduction to UTI and GAMMs will open the paper, followed by a general overview of polar tongue contours GAMMs. Polar GAMMs fitting, significance testing, and model plotting will be illustrated by means of an example pilot study that compares tongue contours of voiceless and voiced stops of 4 Italian speakers. A brief tutorial will show how to fit polar GAMMs with the R package `rticulate`, which has been developed by the author to streamline polar GAMM fitting in R using the `mgcv` package. While the paper focusses on model fitting in R, the overview provided will be relevant to a more general public. Among the advantages of polar GAMMs are higher flexibility in model specification, inclusion of random effects, and control over autocorrelation in the residuals. The paper will conclude by discussing the limitations of the current implementation of polar GAMMs (such as across-speaker normalisation) and its future directions.

References

Davidson, Lisa. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America* 120(1). 407–415. doi:10.1121/1.2205133.

Heyne, Matthias & Donald Derrick. 2015. Using a radial ultrasound probe's virtual origin to compute midsagittal smoothing splines in polar coordinates. *The Journal of the Acoustical Society of America* 138(6). EL509–EL514. doi:10.1121/1.4937168.

Mielke, Jeff. 2015. An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons. *The Journal of the Acoustical Society of America* 137(5). 2858–2869. doi:10.1121/1.4919346.

Wood, Simon. 2006. *Generalized additive models: An introduction with R.* CRC Press.