Towards Automatic Classification of Manner of Articulation and Voicing in French Uvular Rhotics

Andrés Felipe Lara, Claire Pillot-Loiseau
Laboratoire de Phonétique et Phonologie, CNRS & Université Sorbonne Nouvelle

Developing an automatic classification model for French rhotics by manner of articulation and voicing is challenging due to their production variability. French /ʁ/ encompasses approximants, trills, and fricatives, each with voiced and voiceless counterparts [1, 2]. Gendrot et al. [3] described /ʁ/ as a binary continuum between a voiceless fricative and a voiced approximant. Excluding trills risks misclassification, as dorsal fricatives can produce uvular vibrations, creating signals with mixed periodic and aperiodic sources [4]. Perceptual overlaps further complicate classification [5], highlighting the need for robust methods to model rhotic variability.

To address this, we developed a semi-supervised random forest model [6] for multi-class manner and binary voicing classification, using acoustic features: duration, HNR, F0, spectral moments, voicing percentage, and formant frequencies. SMOTE addressed class imbalance by generating synthetic samples for minority classes [7]. Training data comprised 186 tokens (116 fricatives, 45 trills, 25 approximants) from three female speakers of Parisian French in various phonetic contexts. Manual segmentation and classification by two phoneticians resolved previous failed attempts at model training due to F0 estimation errors.

The model achieved 65% accuracy and an F1-score of 0.66 for manner classification, performing well for fricatives (precision 85%, recall 78%) but poorly for approximants and trills (F1-scores 0.44 and 0.37). Voicing classification performed better, with 93% accuracy (unvoiced: precision 97%, recall 90; voiced: precision 89%, recall 96%).

The variability of French /ʁ/, stemming from complex articulatory strategies and overlapping acoustic signatures, remains a challenge. Acoustic features employed fail to capture critical articulatory details. The limited dataset and few speakers restrict generalizability. While SMOTE mitigates imbalance, synthetic samples risk overfitting.

Future work will expand the dataset through additional manual annotation and integrate electromagnetic articulography with acoustic features to gain insights into manner of articulation. Modeling temporal transitions within rhotic segments may further improve understanding of variability across speakers. A probabilistic classification approach is proposed to more accurately capture the variability and diversity of articulatory manners observed. These challenges suggest that strict categorical distinctions for manner of articulation fail to capture the phonetic complexity of French /R/, which appears to exist on a continuum more intricate than previously described [3].

[1] Meunier, C. (1994). Les groupes de consonnes : problématique de la segmentation et variabilité acoustique. *Thèse de Doctorat, Université d'Aix-en-Provence*.
[2] Ramasse, D. (2017). Realizations of the French "r" and their distribution in a word: A preliminary study. *The Phonetician: Journal of the International Society of Phonetic Sciences, 114*, 24–45.
[3] Gendrot, C., Kühnert, B., & Demolin, D. (2015). Aerodynamic, articulatory and acoustic realization of French /ʁ/. *International Congress of Phonetic Sciences*.
[4] Redmond, M., & Jongman, A. (2018). The acoustics of dorsal fricatives: Evidence from production and perception. *Journal of Phonetics, 68*, 1–17.
[5] Solé, M. (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics, 30*, 655–688.
[6] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.
[7] Ganganwar, V. (2012). An overview of classification algorithms for imbalanced datasets. *International Journal of Emerging Technology and Advanced Engineering, 2*(4), 42–47.